State of Wisconsin

CORRESPONDENCE/MEMORANDUM

DATE: December 30, 2003 FILE REF:

TO: Jamie Dunn – NOR/Spooner

DRAFT

FROM: Tom Janisch – RR/3

SUBJECT: Comments On 1) New Fields (NFS) December 15, 2003 Technical Letter Report (TLR) related to Surface Water and Sediment Issues.

General

The discussion below gives my perspective on the main issues at hand and what should be done about the them to get on with the process without any further undue delay. I am admittedly not attuned to the AOC/SOW process between EPA and Excel and what the implications are to the reaching agreements between the NewFields and SEH Work Plans for the proposed round of iterative studies. Also, there are three entities involved in making comments and generating documents - Xcel, URS, and NewFields. In my comments below, I attribute certain statements etc. to the source as I can bet associate them with and may not be correct in this in all regards.

In order to put the NewFields TLR content into some context with what has previously transpired by way of meetings and documents, I also reviewed the following:

1) the Draft August 2002 URS RI/FS Work Plan as it Relates to Surface Water and Sediment Issues.
2) the October 31, 2003 SEH RI Work Plan as it Relates to Surface Water and Sediment Issues.
3) My notes from meetings of:
    a) The EPA Contaminated Sediment Technical Assistance Group (CSTAG) in Ashland of July 15-17, 2002 which included stakeholder presentations;
    b) January 22, 2003 meeting in Madison between WDNR, Excel, EPA, and URS that dealt with problem formulation/assessment endpoint discussions; and
    c) March 26, 2003 meeting in Ashland between WDNR, Excel, EPA and URS that again dealt with problem formulation issues and URS's "Strawman" Baseline Baseline Problem Formulation document.
4) Correspondence /Memorandum and technical reports that I generated as a follow-up to the above Meetings and "Strawman" document. (e.g., January 27, 2002, February 10, 2003, and March 24, 2002.

Based on a review of all the above, the focus issues as I see them are:

1) the ERA problem formulation process as embodied in Steps 1 and 3 of the Superfund eight-step ERA process; use existing problem formulation information or initiate a new problem formulation process?;
2) the design of the supplementary field and laboratory studies contained in the SEH October 31, 2003 RI Work Plan;
3) the use and integration of the existing data collected by WDNR and used in the 1998 ERA and 2002 Supplementary ERA, and integration with this information with the data to be collected under the current plans and designs;
4) The need for *a priori* decision criteria as it applies to 1) weighting the results of the lines of evidence for use in the risk characterization, and 2) deciding what remedial alternative(s) to select based on the risk characterization outcomes.

As part of the review, I have taken the URS comparison tables of the RI/FS Work Plans from their LTR for the Chequamegon Bay sediments and added my comments to them, which incorporates the comments made in the text below. These tables with the comments are in Attachment 4 below at the very end of this document.

Observations/Summary

- Excel/Newfields in their TLR appear to be indicating that the March 2003 URS "Strawman" document should somehow be playing a role in current considerations related to the Problem Formulation component considerations and proposed study designs. My notes from the March 26, 2003 meeting in Ashland where some "Strawman" components were discussed do not identify any clear role or agreements for how the "Strawman" or its contents would be used in future considerations in the iterative risk assessment process for the site. I have no follow up correspondence in my file from the meeting that clarifies this.
- I'm somewhat confused about the Problem Formulation process and Problem Formulation product. My impression is that Excel/URS wants to initiate a whole new Problem Formulation process and ignore the Problem Formulation product that already exists based on the 1998 and 2002 ecological risk assessment studies. A large part of the "Strawman" Problem Formulation document is redundant of information in the existing Problem Formulation product. In relationship to the existing Problem Formulation product, the iterative risk assessment process encompasses such activities as 1) adding new information to the components of the existing Problem Formulation, 2) collecting additional information to support the characterizations and conclusions reached in initial assessments , 3) refining the conceptual site model through new information or studies, 4) re-evaluating site assumptions, and 5) undertaking studies, reviews, or reassessments in response to stake holder and interested party input which is part of the Problem Formulation process. The 2002 Supplementary ERA and the presently proposed studies in the SEH Work Plan are a response to the latter. Xcel/URS appear to have identified some additional information needs they believe should be included for additional studies, some of which need more specificity for consideration in the iterative process. There is no need to redo or reinitiate the Problem Formulation Process each time iterative information is needed or available, but to incorporate any additional information needs into the iterative process and integrate the results with past collected information to arrive at risk characterizations. The question in using this process is how much information is enough and how much iteration has to be done in order to make management decisions for a site within the bounds of accepted uncertainty. We can't quantify the absolute uncertainty associated with each measurement that is made. Such an assessment is not possible. The best that can be done is to convey to risk managers a correct treatment and understanding of uncertainty associated with the lines of evidence used to characterize risks to provide information and insight useful for them in making management decisions in regard to actions. EPA's Principle #5 for managing contaminated sediment risks indicates that the risk assessment framework should not be used to delay a decision at a site if sufficient information is available to make an informed decision. Risk Management Principle #5 states that the risk assessment process should be used to supplement decision making for a site, not supplant it. While we need to follow process, we cannot get hung up in it to the degree that the purpose of using the process to arrive at the end result of risk characterizations are lost or obscured.

- In their comment documents Xcel/URS/NewFields attempt to introduce into the risk assessment process the concept of developing and applying *a priori* or "front end" decision criteria as to 1) how much weight should be given to the various lines of evidence and associated results in the assessment, and 2) how various risk assessment outcomes should be associated with management decisions for the selection of particular remedial alternatives. MADEP (1995), Chapman (1990), Step #5 in the EPA DQO process, and Grapentine et al. (2002) are cited as possible examples to be used in the development of these *a priori* decision criteria. Some discussions in the "Strawman" (e.g., pages 6 and 38) on this issue such as the use of the DQO process above and beyond its normal application of assuring the quality of the data collected was not understandable to me. The process of selecting remedies is a risk management function and not a risk assessment function so any decision criteria use in this regard would take place in the FS process, not the risk assessment process. The Xcel comments on page 9 of the TLR fault the SEH Problem Formulation for making no effort to *"…develop specific 'cause and effect' decision criteria and a specific weight of evidence framework to which all parties have had an opportunity to contribute."* The problem with the Xcel/URS statement above is 1) the EPA risk assessment guidance contains no discussions or guidance on the particular issue of development of such *a priori* decision criteria, and 2) WDNR has not agreed to nor sees a role for such decision criteria in the risk assessment process, so the SEH Work Plan was under no direction to develop and include such criteria. Comments are made below and in the Attachment 2 in regard to these type of decision criteria. An argument can be made that the EPA risk assessment guidance infers or states such *a priori* WOE approaches should not be used in the risk assessment process because of their limiting and inflexible nature. The literature is replete with various WOE schemes but there is no agreement on a definitive approach nor are there attempts to relate such schemes to the EPA risk assessment guidance. Attempts at discussions to derive such *a priori* decision criteria could be made by the parties but my recommendation would be not to do so and follow the risk assessment guidance that does not identify a role for such criteria in the process. If such criteria are to be discussed, it is incumbent on Xcel/URS to come up with a draft of a detailed framework and decision criteria they want to see used and not the general references they have used to date. Also, assuming that they have seen the benefits of the development and application of such decision criteria at other Superfund contaminated sediment sites, they should be able to provide examples where management decisions have been made at these sites based on these types of decision criteria. Their statements on why these type of decision criteria should be developed and applied are noticeably absent in this regard despite their statement that *"Recent EPA and other federal agency guidance is replete with this advice."* One question is what particular EPA guidance are they referring to in this regard? Arguments about the data regarding quality, adequacy, and interpretation will likely not be resolved by such attempts to arrive at *a priori* decision criteria any more than the arguments at the end of the process are likely to be.

- It was also noted in the TLR that Xcel faulted the SEH Work Plan for purportedly making no effort in their Problem Formulation to consider EPA's eleven principles for managing contaminated sediments. This issue is commented on below. As noted, WDNR has responded to the CSTAG group's recommendations on each of these principles and those that can be translated into actionable items were. Some of the principles such as numbers 9, 10, and 11 have more to do with remedy selection and would not be involved in the risk assessment process. Excel/URS presents few specifics on how they think the principles should be translated into actionable items in the Problem Formulation process. While their Draft RI/FS Work Plan makes a general statement that a *"more formal and systematic integration of concepts introduced in the guidance for managing contaminated sediment sites (USEPA 2002) into the risk assessment process.",* they provide no specifics as to how this should be done, and I see nothing in their Work Plan or Problem Formulation efforts where they have identified how they intend to integrate each of the principles into their proposed efforts.
- At this point in time, I would have thought Xcel/URS would have moved from repeating their more general comments and critiques to putting more specifics and details on the table for discussion for their proposals by means of examples, needed study components, study designs, changes to SEH study designs, etc. to advance the risk assessment iterative process to reach a final risk assessment product upon which decisions can be made. The TLR does little in this regard.

1) Problem Formulation Process Issues

In all of the URS/NFS documents and comments there seems to be the continuing reference to the need to come up with or draft a new Problem Formulation document that would seem to be related to initiating a new Ecological Risk Assessment process for the site. These comments overlook the fact that an ERA and a supplemental ERA have been completed for the site and contain the attendant information and data that constitutes the base Problem Formulation for the site. URS/NFS have stated that a "formal" Problem Formulation is required to support both a Baseline Human Health Risk Assessment and Baseline Ecological Risk Assessment. URS/NFS repeatedly state in their "Strawman", their August 2003 RI Work Plan, and in their TLR that "a more formal and systematic integration of concepts" contained in EPA's (2000) eleven *Principles for Managing Contaminated Sediment Risks at Hazardous Waste Sites* into the risk assessment process and principally the Problem Formulation process for the Ashland site. They believe synthesis of this guidance will somehow allow the "systematic and objective basis" for addressing issues in OU-4 of the Site. NFS contends (page 9 of TLR) that SEH has made no effort to consider any of the eleven principles in their Problem Formulation (other than reference to a sediment stability model) [note, as discussed below It SEH has put forward study designs for field and laboratory that are intended to supplement the exisiting Problem Formulation information, not write a new Problem Formulation]. It is assumed that wanting to address the eleven principles are the basis for Excel/URS wanting to redo/restart the Problem Formulation process from scratch for the Site and ignoring the Problem Formulation base product that exists. Also, we note that URS has not identified where they have made an effort to consider the eleven principles for managing contaminated sediment in their Problem Formulation. Where are these principles discussed in Problem Formulation in their work plan?

NewFields in their TLR make several references the "Baseline Problem Formulation process that was agreed to by all parties" in the March 2003 meeting in Ashland. NewFields also state that "Xcel Energy proposes that the Baseline Problem Formulation process be initiated as was originally intended using the SEH workplan and the URS 'Strawman' as the bases." I have no documents in my file that summarize the March 2003 meeting discussions or expectations of the participants. My notes from the March 2003 meeting reflect the establishment of a very tentative schedule related to the SEH RI Work Plan, which we now have in hand, and discussions of some of the specifics of the URS "Strawman".

From my perspective, the Problem Formulation Process and the resulting document product can be an evolving and supplemented document, as the particular situations of a site require. As hypotheses and conclusions are tested, as applicable new research literature becomes available, as re-valuations of site assumptions and refinements of the conceptual site model are done, as new information about the site become available, and as data gaps in site information and studies are identified, decisions are made as to how to supplement components of the existing Problem Formulation document. The supplementation takes the form of an iterative process that results in adding information to the existing Problem Formulation document. The emphasis is on iterative and revising, not reinitiating the process, in other words not scrapping an existing Problem Formulation document and restarting the entire process as Excel and their consultants seem to be advocating. Following the above as it relates to the iterative process and supplementing the existing Problem Formulation document as needed, there is no need to initiate a completely new Problem Formulation process every time a component is supplemented, which would represent a waste of already expended time and money spent on the process. Large portions of the March 2003 "Strawman" Baseline Problem Formulation document represent a redundancy of portions of the contents of the Problem Formulation components/content in the existing documents and sections of the existing Ecological Risk Assessment for OU-4 as an initial effort to restart and redo the whole process. I never really understood what the purposes of all this repetition was and the rationales for redoing an already completed process. It would seem the focus on the part of URS/Excel should have been a listing of specific components in the existing Problem Formulation they were not in agreement with and those components they wanted to see supplemented in the iterative risk assessment process..

URS/Excel on page 3 of their “Strawman” seem to recognize that the risk assessment process can be an iterative process and can benefit from work and planning done previously. However based on their approach in their “Strawman”, they are not recognizing the fact that there is an existing Problem Formulation component for the site. URS/Excel also recognize EPA risk management Principle #5 which is *Use an Iterative Approach in a Risk-Based Framework*. However, as discussed above their approach seems more repetitive or redundant, than iterative. Some of the concepts contained in EPA’s Risk Principle #5 are as follow:

1) The risk assessment framework is intended to supplement, not supplant, the CERCLA remedial process mandated by law for Superfund sites.
2) Although there is no universally accepted, well-defined risk-based framework or strategy for remedy evaluation as sediment sites, there is widespread agreement that risk assessment should play a critical role in evaluating options for sediment remediation.
3) EPA encourages the use of an iterative approach, especially at complex contaminated sediment sites. Each iteration might provide additional certainty and information to support further risk-management decisions, or it might require a course correction.
4) An iterative approach may also incorporate the use of phased, early, or interim actions at some sites to reduce risks or to control ongoing spread of contamination.
5) Citing the NRC report: “The [NRC] committee cautions that the use of the [risk assessment] framework…should not be used to delay a decision at a site if sufficient information is available to make an informed decision. Particularly in situations in which there are immediate risks to human health of the ecosystem, waiting until more information is gathered might result in more harm than making a preliminary decision in the absence of a complete set of information. The committee emphasizes that a ‘wait-and-see’ or ‘do-nothing’ approach might result in additional or different risks at a site.”

URS state on page 4-1 of their August 2003 RI/FS Work Plan *that “the agencies recognize that the previous risk assessment studies are deficient…and a formal Problem Formulation particularly with regard to OU-4 are required to support the Baseline Risk Assessments.”* As part of the risk assessment process, review of the risk assessment products have garnered comments that in order to address, have necessitated additional iterative studies. The results of these studies have resulted and will result in supplementing the existing Problem Formulation contents of the ERA. They will not necessitate initiating or redoing a new Problem Formulation process. URS is incorrect in stating this agency feels a new Problem Formulation process needs to be initiated or the Problem Formulation process needs to be redone. As is part of the iteration process done in risk assessments, the existing Problem Formulation product will be supplemented or revised as necessary, but not redone. URS is generally wrong on both issues in their above statement.

NewFields state on page 8 of their TLR that “SEH has performed biological sampling and developed two separate preliminary ecological risk assessment reports.” NewFields is incorrect in the characterization of these documents. The initial ERA report was a completed ERA, not a “preliminary” document. The use of the word “preliminary” with the ecological impacts characterized from these ERA as in Section 3.4 of the SEH Work Plan is related to the standard process in the risk assessment that calls for the derivation of preliminary remediation goals (PRGS) from the data. Following the iterative process discussed above, the second ERA was a Supplementary ERA that incorporated new information collected to respond to comments on the original document. It was a supplemented ERA and resulted in supplemented information to the existing Problem Formulation contents. The ERAs were not separate and neither were they preliminary, so NewFields is wrong on both counts in attempting to characterize these documents.

It should be noted that the current studies designed by SEH and contained in their October 2003 RI Work Plan are a part of the iterative risk assessment process and are the principal result of responding to the

CSTAG and stakeholder inputs presented during the CSTAG meetings over July 15-17, 2002. The design of and results of the studies and supplementing the conceptual site model are for the purposes of responding to these inputs and will be used to supplement the existing ERA and Problem Formulation process in an iterative and integrative manner. The studies will not be used for a stand alone ERA as the comment made in the NewField's TLR implies on the bottom of page 9. NewFields is associating the number of samples involved in the SEH Work Plan with SEH's *"…entire risk assessment analysis program."* NewFields is incorrect in this regard. The final risk characterization process will be based on a combination of the results of all samples collected in the original ERA, the first iterative studies that resulted in the Supplementary ERA, and second round of iterative studies what will be incorporated into a second Supplementary ERA. Further, these sample results will be integrated with all the other lines of evidence for risk characterization purposes and no single line of evidence will constitute the *"entire risk analysis program"* as NewFields seems to imply. See Section 4.2 of the SEH RI Work Plan in this regard.

On the issue of the EPA's eleven principles for managing contaminated sediment risks, it should be noted that WDNR has responded in an Oct. 16, 2002 letter to EPA to each of the September 3, 2002 CSTAG site-specific recommendations related to each of the eleven principles. WDNR has identified how each of the site-specific recommendations will be addressed. It is unclear to me how NewFields is now transitioning these eleven principles to the SEH Work Plan by stating on page 9 of the TLR that *"There appears to be no effort in the SEH Problem Formulation to consider these principles (although a reference to developing a potential sediment stability model in Section 5.5.3)."* SEH has designed studies as part of the iterative risk assessment process related to the existing Problem Formulation but is not redoing the entire Problem Formulation process so it is unclear to me how NewFields is relating the eleven principles to the Problem Formulation process in this case. I am not aware of any specific EPA guidance for incorporating the eleven generic principles into specific actionable items in the Problem Formulation process and even if they all need to be translated into actionable items in all cases. The EPA eleven principles are guidance for project managers to follow in evaluating cleanup alternatives, and establishes a consultation and review process for remedial and removal actions. The sediment principles provide interim policy guidance while EPA continues its work on a more detailed technical manual, *Contaminated Sediment Remediation Guidance for Hazardous Waste Sites* (Goodwin Proctor LLP. 2002. To dredge of not to dredge? Risk determines remedy under EPA's sediment Principles. Environmental Law Advisory. March 2002).

Problem Formulation and Addressing Sediment Stability

It appears the one principal that URS/NewFields is focusing on related to their emphasizing the eleven principles for incorporation into some part of the risk assessment process is principle #4 that states *Develop and Refine a Conceptual Site Model that considers Sediment Stability.* However, addressing this principle does not require a new Problem Formulation process. The appropriate place for inclusion and addressing the sediment stability question is in the existing Conceptual Site Model that was formulated as part of the Problem Formulation and its iterations. The NewFields questions related to sediment stability is how does the existing Conceptual Site Model need to be revised to address the sediment stability issue as posed by URS/NewFields.

The proposed tasks for the preliminary evaluation of sediment stability originally introduced in the March 2003 "Strawman" and discussed in the URS August 2003 Draft RI/FS Work Plan seem to involve a multi-phased effort with the screening level phase transitioning into detailed modeling effort if necessary. The proposal involves complicated modeling that would likely render any modeling predictions highly variable and uncertain. The two possible conceptual models related to backfilling source and possible watershed runoff at the time of the logging era seem hard to relate to the present state of contamination in the bay. Will only these two conceptual models be used to attempt to explain the current contaminant distribution at the site? What if one or both of the possible conceptual models that URS is using to explain the current contamination distribution at the site are incorrect premises for trying to explain the current contamination? Is it normal to limit the basis for the modeling to only such a limited number of preconceived conceptual models? What use with the modeling results be if it is based on incorrect assumptions at the starting point? What are the timelines for completing this modeling effort including the detailed modeling if necessary so the information can be used in the management decisions for the site. One of my questions is why is it important to attempt to describe the historical development of the existing contaminated sediments by considering various possible sediment and contamination transport mechanisms and pathways? The proposal also indicates that the conceptual model may also be used to estimate historic site-specific transport rates and provide a basis for estimating future rates of sediment erosion and deposition. Of all the sediment stability issues to be focused on for the site, the latter as it relates for estimating future rates of sediment erosion, deposition, and contaminant transport would seem to be the most important from the management decision standpoint and should be prioritized in any model development effort. Also, the stability issue should be based not only on the sediment stability issue alone but the broader substrates types in the bay that includes contaminated wood waste materials.

Given the present degree of contamination in the bottom of the bay, its likely decades long presence, and the limited potentialities for attenuation by other than dispersion (which is not acceptable) based on the discussions in the ERA for the site, it appears the primary management decisions for the site should be based on the outcomes of the present and future risks risk characterizations. The latter assumes the bottom conditions of the site will not change in future decades from the past decades. To prevent any more delays for making management decisions for the site, any modeling efforts need to be completed in a time line that closely follows when the risk characterization process is completed for the site and risk assessors will be providing this information to the risk managers for the site for decision purposes.

2) Design of the supplementary field and laboratory studies contained in the SEH October 31, 2003 RI Work Plan

Xcel Energy seems to be making conflicting statements in their TLR in regards to their agreement/disagreements with the studies proposed in the SEH Work Plan. On page 9, they state that the testing programs proposed in the SEH Work Plan appear to be appropriate (except for the number of samples) but on page 10 they state Xcel Energy believes it is inappropriate to critique the specific studies described by SEH in the Excel TLR because they feel this has not proven a productive process in the past. They also state that they are proposing that the Baseline Problem Formulation process be initiated as was originally intended using the SEH work plan and the URS "Strawman" as the basis. If the Baseline Problem Formulation process Xcel is referring to is the input from all parties and stakeholders in the iterative risk assessment process, then the process is for Xcel to provide comments that are specific to the details of the study designs in the SEH Work Plan or state that they want additional studies, and provide details and rationales for those studies, whether they originate from the "Strawman" or elsewhere. This is the process and there is no need to "initiate" it. To advance the status of the studies, Excel needs to provide their detailed comments and critiques of the SEH designed studies and provide for any additional studies in detail to serve as points of discussion.

Xcel goes on to say that to initiate this process (Baseline Problem Formulation process) that they have incorporated all existing sediment into a GIS platform. The contamination profiles and isopleths will serve as a tool of reference for principally locating sites to be representative of the gradient of surficial substrate contaminant concentrations over the site to be used in the planned iterative studies. Related to the selection of these sites, Xcel makes a somewhat confusing statement on page 9 of the TLR when they state *"the potential exists that higher zones of contamination near one of the proposed sampling locations will have adverse effect on the resulting data, leading to biased conclusions.* Selection of representative contaminant concentrations across a gradient of concentrations over the site from high to low means representative high end sites must be included. How these sites should not be excluded as Excel seems to be saying will bias the results low in that no endpoint effects information from the high end of possible exposure concentrations will be included in the data set to be used in the risk characterizations. This does not make sense.

It is also confusing that URS is stating on page 8 of the TLR that their work plan describes *"a number of other validation studies (Section 5.1.4.5) that would potentially be developed pending the conclusions of the Problem Formulation process. It was URSs understanding at that time that, consistent with ERAGs guidance and contaminated sediment management principles, completion of this process would involve agency representatives as well as other Interested Parties."* A review of Section 5.1.4.5 in the URS Work plan and the comparison table at the end of the TLR for the most part list the studies already proposed in the SEH Work Plan. The studies proposed by SEH as discussed above are a result of the iterative risk assessment process and are largely in response to comments made at the July 2002 CSTAG meeting by EPA, various stakeholders, and interested parties. The process by which this has been done is consistent with risk assessment guidance and there is no reason to repeat it. The next step in the iterative process would be to provide the results of the studies to the parties in the form of the supplemented ERA or HHRA. The studies are being done in the context of the existing Problem Formulation product and process and there will be no waiting for potential development of studies pending the conclusions of the Problem Formulation process.

URS has compared the URS conceptual plan to SEH's more detailed work plan and indicated the comparisons were difficult because URS's work plan was conceptual and tentative in nature. However, to advance this process, what are needed are specificity and details about any proposals and not concepts and tentative ideas about what studies URS believes are needed. While I get some idea of the questions URS has about some of the SEH study designs, I'm not left with a clear idea of how much differences there are in other study components because of the conceptual and tentative nature of URS's proposals. Again, more specificity is needed about any questions URS has on the SEH study details.

Number and Selection of Sample Sites

One of the components of the SEH study designs that URS questions is the number of sites to be sampled. Other than stating they don't agree with the number, they do not provide a number of samples that would be satisfactory to them and provide the rationales for that number of samples. This is an example of where specificity in terms of providing a number of samples satisfactory to them with the supporting rationales would provide for a discussion point to get the issue resolved and get the studies underway in a timely manner. Responses to increase the proposed number of samples on SEH's part without this specificity on URS's part would likely result in another response about the dissatisfaction about sample numbers. This will not advance the issue. With URS now apparently undertaking the responsibility of conducting the studies, the number of samples they collect will need to be satisfactory to them and based on their comments it will be some number greater than 8.

In the comparison table of the LTR, URS has some questions in regard to how the off-site reference sites will be selected based on the SEH Work Plan. Clarification in this regard may be needed. It is assumed that the same design components used in the original ERA that assessed coal tar contaminants in the separate wood substrate separate from the silty sand substrates still applies. In other words, enough samples need to be taken in each type of substrates and over a representational range of contaminant concentrations in both to do the assessment. This will also require selection of representative background sites for each type of substrate. I commented on the issue of the adequacy of the number of representational study sites from each substrate (3) and the background sites in my Dec. 24, 2003 comments on the some draft SEH study designs. The whole issue of number and locations of study site samples and number and location of reference site samples is a design component that will apparently need to be revisited and agreements reached between the parties.

URS Pore Water Characterization Study Proposals

One of the possible biological studies URS has listed Secion 5.1.4.5 of their August Work Plan and in their TLR comparison table is for pore water characterization. URS first brought up pore water-related studies in their March 2003 “Strawman”. URS notes that SEH has not included pore water characterization in their study designs. Again to advance the issue, URS should have developed a tentative study design to use as point of discussion/decisions containing specifics on what type of characterizations would be done on the pore water (e.g, chemical, physical, and/or biological), what are the testable hypothesis, how the pore water would be collected, what procedures will be used to minimize changes to the *in situ* condition of the water (most sediment collection and processing methods have been shown to alter interstitial water chemistry thereby altering contaminant bioavailability and toxicity [EPA, 2001]), what depth of sediment will be collected for the pore water extractions, the advantages and disadvantages of various pore water extraction methods, etc., and of doing pore characterizations in general. Decisions on whether or not to include pore water testing in the proposed studies at this point in time will be limited by not having the above information in hand. URS had almost a year to develop this type of needed background information for pore water characterization studies. If URS felt strongly about including pore water characterization as an iterative study component in the risk assessment they should have this type of information available by now.

US EPA. 2001. Methods for collection, storage, and manipulation of sediments for chemical and toxicological analyses: Technical manual. EPA-823-B-01-002. October 2001.

URS Wood Waste Study Proposals

Section 5.1.4.6 of the TLR and the comparison table in this document propose conceptual studies for evaluation of wood waste impact. These proposed study on wood waste impact were contained in the March 2003 “Strawman”. The proposed study appears to go beyond the studies conducted in the past ERAs that attempt to distinguish the impacts from coal tar contaminated wood substrates from uncontaminated wood substrates. URS is proposing to study the impacts of wood waste substrates on the benthic communities compared to communities that are expected to be present on the uncontaminated natural mineral substrates on the bay bottom. As discussed in my previous comment paper (Attachment 1), dealing the presence of wood waste materials as anthropogenic physical stressors would appear to go beyond the scope of the Superfund program set up to address released chemical stressors into the environment.

3) The use and integration of the existing data collected by WDNR and used in the 1998 ERA and 2002 Supplementary ERA, and integration with this information with the data to be collected under the current plans and designs.

Nothing in the URS “Strawman”. Draft RI/FS Work Plan, or TLR gives an indication of how URS intends to integrate the data from the WDNR 1998 ERA and 2002 Supplementary ERA with the data that will be collected under the Work Plan being reviewed to produce a second supplemented ERA. The above documents appear to indicate Excel/URS are planning to initiate a new stand alone ERA based on the results of the studies under review. An indication of this is on page 9 of the LTR where they are wrongly assuming the *“SEH work plan appears to base its entire ecological risk assessment analyses program on the results of data from eight sample locations.”* Sections 5.6.2 and 5.6.2.2.1 of the SEH work plan clearly indicate that the all the data including the data in the original ERA and the data collected in the two iterative studies will be used be used do the risk assessment and risk characterizations for the Site. Excel/URS needs to provide information in their work plan on how they intend to handle and integrate all the data from the three studies. An example of an approach to integrate data from the toxicity testing from all the studies to arrive at preliminary remediation goals (PRGs) for the contaminated wood and silty sand substrates is shown in Attachment 3 below. The evaluation was done as a result of the CSTAG comments and provides another approach to the one used in the 2002 Supplementary ERA to derive the PRG values. This alternative approach arrives at similar PRGs to those in the Supplementary ERA and supports those PRG values. The approach is transparent and is based on protecting 80% or greater of the benthic organisms related to differences in endpoints effects compared to the reference site results. The ranges of contaminant concentrations related to levels of protection derived from an approach such as this can be used by risk managers in their decision making process.

4) The need for decision criteria as it applies to 1) weighting the results of the lines of evidence for use in the risk characterization, and 2) deciding what remedial alternative(s) to select based on the risk characterization outcomes.

Originally contained in the March 2003 "Strawman", Excel believes it is important to establish *a priori* or on the front end of the process, decision criteria that would be related to each of the decisions points above. The former is related to a risk assessor decision point as to the use of data, and the latter is related to a risk manager decision point as to possible remedy selection. Proposals or suggestions on how to develop these decision criteria are contained in Sections 3.4.3, 3.5 and 3.6 of the "Strawman".

Statements on page 9 of the TLR in this regard are:

> *" URS' Strawman report recognized that in addition to the proper development of this decision making process…"* and *"In addition, there appears to be no attempt [in the SEH Work Plan] to develop specific 'cause and effect' decision criteria and a specific weight of evidence framework to which all parties had an opportunity to contribute. As a result, the process is vulnerable to the same shortcomings that have already been experienced. Namely, if this program is performed in accordance with this plan the parties will debate how the outcome of the studies will be used to support risk management decisions."*

Some observations on the above in regard to establishing *a priori* decision frameworks for utilizing the lines of evidence and in remedy selection:

- URS is stating the DQO process should be used develop an *a priori* decision framework that considers which remedies will be implemented based upon a range of risk assessment results (Section 3.4.3, page 38, Strawman).
- The URS RI/FS Work Plan (page 4-1) state that the Data Quality Objective (DQO) process is described in USEPA guidance is a *"seven-step planning approach to develop sampling designs for data collection activities that support decision making. It is recommended by USEPA in…ecological risk assessment guidance (USEPA 1997; 1998)."*
- It should be noted that USEPA risk assessment guidance distinguishes between the use of the DQO process in analysis planning in the Problem Formulation stage of the risk assessments and how the DQO process is used in other applications. This is clarified in the following taken from the USEPA (1998) risk assessment guidance:

  > *"Analysis planning is similar to the data quality objectives (DQO) process which emphasizes identifying the problem by establishing study boundaries and determining necessary data quality, quantity, and applicability to the problem being evaluated. The most important difference between problem formulation and the DQO process is the presence of a decision rule in a DQO [Step 5. Develop decision rule] that defines a benchmark for a management decision before the risk assessment is completed. While this approach is sometimes appropriate, only certain kinds of risk assessments are based on benchmark decisions. Presentation of stressor-response curves with uncertainty bounds will be more appropriate than…decision criteria where risk managers must evaluate the range of stressor effects to which they compare a range of possible management options (see Suter, 1996)."*

- The 1997 USEPA guidance that contains the process for designing and conducting ecological risk assessments discusses the use of the DQO process but does not include Step 5 of the DQO process related to establishing defining benchmarks or decision criteria for selection of remedies. Based on the above USEPA guidance, it would not appear appropriate to use Step 5 in the DQO process to develop an *a priori* decision framework that considers which remedies will be implemented based on a range of risk assessment results. Site specific factors and characteristics, management goals and objectives, unknown study outcomes, and the need to be flexible in making management decisions on a site-specific basis would make it difficult to anticipate and tailor an a priori decision framework for making management decisions related to remedy selection.
- Excel/URS also want a weight of evidence decision framework that will be used to "weigh" the individual lines of evidence for importance before they are used in the risk characterization process. My perspective on developing weight-of-evidence approaches in an *a priori* manner for application to the lines of evidence before doing the data integration and characterizing risks is in Attachment 2 below. In my opinion, WOE approaches that involves attempting to quantitatively or qualitatively assign weights to lines of evidences have little or no role in risk assessments. Uncertainties associated with lines of evidence are identified and discussed in the uncertainty analysis section of which are a standard part the risk assessments documents.
- The USEPA risk assessment guidance contains no discussion of the use of WOE approaches as part of the risk assessment process. What the guidance has to say about using and integrating lines of evidence in a strength of evidence approach is in my attached discussion paper below (Attachment 2).
- While URS has provided generalities on the development of decision criteria and decision frameworks in their "Strawman", they have provided no specifics or examples developed by them to use as a starting point of any discussions on the matter.
- More importantly, it would have been useful if Excel/URS could have provided examples of these decision frameworks that have been developed and applied at other Superfund sites with contaminated sediments to show where management decisions were made at the sites based on developed decision criteria and decision frameworks.
- Since the WDNR does not see a role for decision frameworks or decision criteria of these types, SEH was not instructed to develop or include them in the Work Plan they developed.

ATTACHMENT 1

Perspectives On and Context Of Restoration and Protection of the Nearshore Habitat Off Kreher Park. Addressing Wood Waste Materials on the Bottom of the Bay.

The shoreline on Chequamegon Bay adjacent to the City of Ashland has undergone a number of alterations since establishment of the city in the late 1880’s up to the present day related to commercial, industrial, and urban developments and uses. The alterations include construction of jetties, ore docks, other filled dock areas, and placement of various types of fill materials along the shoreline. Building construction and asphalt and concrete paved areas have decreased the land areas for infiltration and results in nonpoint source discharges to the shoreline areas of the bay. At one time or another, various industrial and other point source discharges may have occurred. Restoration and protection of the the present nearshore habitat areas given this history has to beput into the perspective of what can be attained and is practical given what may be the constraints of the present state of urban development. That is, it would be impractical to have as a management goal for the site the restoration of the nearshore area as it was pre-development. At least for the nearshore habitats in urban areas, some degree of possible human-modified biodiversity has to be recognized in the establishment of management goals for the area off Kreher Park.

It has been documented that due to the operation of a series of sawmills on the shoreline of what is now Kreher Park from the 1880’s until about the mid-1930’s, wood waste in the form of saw dust, wood chips, and various sized wood pieces were placed into the nearshore area. Additionally, rafting of logs into the nearshore areas resulted in tree bark materials being deposited into the areas. Over time, these wood waste materials have been distributed throughout the bottom areas of Chequamegon Bay.

The woody materials in the bottom areas provide a different bottom substrate type for benthic invertebrate communities compared to the silty sands and sands substrates found in the bay. The physical characteristics and microhabitats of each type of substrate will have a role in determining the abundance and diversity of benthic macroinvertebrate that will be established on each substrate type. Each community will be based on the organisms that are tolerant of or can adapt to the physical and chemical conditions presented by each substrate type. Results from the 1998 SEH ERA indicated that the diversity and abundance of benthic macroinvertebrates was less in the wood substrate compared to the silty sand substrate.

The design of the 1998 ERA and Supplemental 2001 ERA for the Ashland site under Superfund guidance did not consider the presence of the wood waste as a stressor that needed to be considered or focused on. The focus for the ERA, as for any ERA under Superfund was chemical contamination stressors, in this case the coal tar-related contamination. The 1998 and Supplemental 2001 ERA study designs were based on looking at how the coal tar-related contamination impacted the organisms associated with both the wood and silty sand substrate types and were not designed to look at the wood waste as an additional stressor. The points of comparison for determination of impacts were comparisons of paired chemically impacted and nonimpacted wood and sand substrates. The ERA study design was not based on doing a cumulative risk assessment involving physical stressors related to the wood waste and the chemical coal tar stressors, but only the latter alone based on Superfund guidance.

As discussed above, there have been a number of alterations of the shoreline areas that have introduced a possible number of anthropogenic physical stressors to the nearshore habitats and the organisms in those habitats. The role of an ERA for Superfund is not to deal with these type of anthropogenic physical stressors. The role that these anthropogenic physical stressors play in determining the degree that habitats and organisms can be restored after the Superfund-related chemical stressors have been removed needs to be identified in the management objectives to lend practicality to the restoration and recovery efforts.

If interested party and community input indicates that they feel the wood waste presence in the bottom areas off Kreher Park associated with the site need to be addressed along with the coal tar contamination, they would need to be informed that addressing by removal would need to be done by another program outside of the Superfund program. Also, any removal from a localized area needs to be put into the context of what influence the presence of wood waste outside of the localized area would have over time. Since the bottom areas of Chequamegon Bay are covered by wood waste, over time the wood wastes from these areas could be transported into localized nearshore areas where removals have been done. Under this scenario, it would appear necessary to address the wood waste problem on a broader geographic scale then a localized scale.

It is expected that directly or incidentally some portion of the wood substrate materials over the nearshore area off Kreher Park will be addressed or removed depending on the remediation alternative selected. The Superfund program and the ERA results based on addressing the chemical contamination may call for addressing that portion of the wood substrate that is contaminated by coal tars. Incidental to addressing or removing coal tar contaminated sand and silty sand substrates, it will be necessary to also remove overlying wood materials whether contaminated or not. Under a removal remediation scenario, it is expected that a large portion of the overlying wood substrate materials over the nearshore area off of Kreher Park would be addressed.

ATTACHMENT 2

DRAFT

The Role of A Weight-of-Evidence Approach Based on the EPA Ecological Risk Assessment Process Guidelines

The Weight-of-Evidence (WOE) Approach In the ecological risk assessment is done by evaluating separate lines of evidence which are represented by the individual measurement endpoints selected during the problem formulation step of the risk assessment. The measurement endpoints are related to the assessment endpoint or the actual environmental component or its attributes that are to be protected. Multiple measurement endpoints can be associated with a single assessment endpoint. The data and results of the measurement endpoints are used and evaluated by: 1) comparisons made of the concurrency and/or non-concurrency of the results from the individual measurement endpoints, e.g., are the results from the separate measurement endpoints of the same relative magnitude of difference when compared to the reference site results. 2) evaluating the magnitude of differences in measured results between the study sites and reference site , and 3) integrating the results of the measurement endpoints and used as lines of evidence in the risk estimation along with other information to establish the strength of the evidence to show the likelihood of ecological impacts. The uncertainty analysis of the risk summary identifies those attributes of each line of evidence that may influence the level of confidence in the use of the results from that line of evidence in the overall risk estimate. The results of the WOE approach and uncertainty analysis used in the risk characterization and risk estimate must be conveyed to the risk managers for the site and to the stakeholders in an understandable and coherent manner.

A definition of the "weight-of-evidence approach" from Warren-Hicks and Moore (1998) is"

> "Weight-of-Evidence Approach – The results of an evaluation of multiple lines of evidence in an ecological risk assessment. A weight-of-evidence approach reduces many of the biases and uncertainties associated using only one approach to estimate risk. The lines of evidence that may be considered in a weight-of-evidence approach include comparing levels in the environment to the levels in laboratory bioassays, field observations, in situ tests, ecoepidemiology, and population and ecosystem modeling. Each line of evidence is evaluated for relevance of the evidence to the exposure scenario of interest, relevance of the evidence to the assessment endpoint, confidence in the evidence or risk assessment, and likelihood of causality."

The biggest issue in the WOE approach is how much weight or if any weight or different degrees of importance should be given to an individual measurement endpoint relative to the other measurement endpoints in using them in the characterization of risks to the assessment endpoint being evaluated. Various quantitative and qualitative weighting schemes have been proposed to assign a relative weight or level of confidence to each individual measurement endpoint relative to its contribution in characterizing and estimating risk to the assessment endpoint involved. There is 1) a call for some type of quantitative, formalized weighting method for lines of evidence although it is unclear what this formal weighting is to consist of and how consistencies and inconsistencies from the results of lines of evidence are defined and determined, and 2) that the process be transparent and done early on in the risk assessment process (problem formulation stage) so that all parties will know how the measurement endpoint data will be used to characterize and determine risk.

The weighting schemes are attempts to eliminate what are viewed as the subjective and qualitative nature of how different risk assessors weigh, interpret, and apply the results from various measurement tools used as lines of evidence. Developers of such schemes claim their approaches lend more of an objective and quantitative basis for weighing the lines of evidence but in fact such schemes are mostly grounded in the professional judgements of the risk assessors, which can result in subjective and qualitative results. Menzie et al. (1996) state that *"professional judgements applied in the selection and evaluation of measurements may incorporate both knowledge about the strengths and weaknesses of various measurements and beliefs about whether the measurements in question are likely to overestimate or underestimate risk."* As a result, regulatory agency risk assessors, who are charged with protection against harm may apply and interpret the results of measurements different from risk assessors representing the regulated community based on professional judgment, experience, or familiarity and use of particular measurement endpoints by a risk. Menzie et al. also state *"a formal weight-of-evidence evaluation will not eliminate the influence of such beliefs from professional judgment. It may, though, increase risk assessor's awareness of his/her beliefs, and elucidate for the user/reviewer of the assessment the beliefs on professional judgment."* At best weighting schemes characterize professional judgements and allow others to see it in this light.

Quantitative weighing schemes have also been perceived as being to inflexible to allow for consideration of all pertinent data and information in a risk assessment. Some pertinent, useful information/data may not fit into weighing schemes but it is important to include a careful evaluation of the information to help understand and explain observed results and relationships among the lines of evidence. Also, some line of evidence weighting schemes are viewed as being overly complicated and to mechanistic. A qualitative approach instead of a quantitative approach may be more understandable and useful while still maintaining the process of characterizing professional judgements regarding the attributes of the lines of evidence.

Difficulties can be encountered in attempting to assign quantitative weights to the individual measurement endpoints even though risk assessors for the regulator and the responsible parties may agree that the weighting should be attempted and used in the WOE approach. Professional judgment, experience, or familiarity and use of particular measurement endpoints by a risk assessor may result in assignment of different quantitative weights to those endpoints compared to another risk assessor's values.

There is presently no consensus of exactly what it means or how a weight-of-evidence approach in a risk analysis involving contaminated sediments should be carried out (Menzie et al. 1995). Similarly, there is no consensus of how weights in the form of quantitative values or qualitative descriptors to indicate varying degrees of importance should be derived and assigned to individual lines of evidence. U.S. EPA's Guidelines for Ecological Risk Assessment indicates a preference for using the term "line of evidence" rather than WOE based on the following discussion extracted from their Ecological Risk Assessment Guidance:

> *"For ecological risk assessments that entail more than one type of study (line of evidence), a strength-of-evidence approach is used to integrate different types of data to support a conclusion. The data might include toxicity test results, assessments of existing impacts at a site, or risk calculations comparing exposures estimated for the site with toxicity values from the literature. Balancing and interpreting the different types of data can be a major task and require professional judgement."*
>
> *The development of lines of evidence provides both a process and a framework for reaching a conclusion regarding confidence in the risk estimate. It is not the kind of proof demanded by experimentalist (Fox, 1991), nor is it a rigorous examination of the weights of evidence. (Note the term "weight-of-evidence" is sometimes used in legal discussions or in other documents, e.g…Menzie et al. 1996). The phrase "lines of evidence" is used to de-emphasize the balancing of opposing factors based on the assignment of quantitative values to reach a conclusion about a "weight" on favor of a more inclusive approach which evaluates all available information, even evidence that may be qualitative in nature. It is important that risk assessors provide a thorough representation of all lines of evidence developed in the risk assessment rather than simply reduce*

their interpretation and description of the ecological effects that may result from exposure to stressors to a system of numeric calculations and results.

Confidence in the conclusions of a risk assessment may be increased by using several lines of evidence to interpret and compare risk estimates. These lines of evidence may be derived from different sources or by different techniques relevant to adverse effects on the assessment endpoints, such as [hazard] *quotient estimates, modeling results, or field observational studies.*

There are three principal categories of factors for risk assessors to consider when evaluating lines of evidence: 1) Adequacy and quality of data, 2) degree and type of uncertainty associated with each line of evidence, and 3) relationship of the evidence to the risk assessment questions."

Based on the above WOE approach, the EPA guidelines do not advocate weighting individual lines of evidence and based on this, do not delve into possible weighting schemes or their application in risk assessments.

It is interesting that while consultants insist on strict adherence to the U.S. EPA guidelines for conducting ecological risk assessments, they want to introduce WOE approaches that call for *a priori* establishment of ranking and weighting of measurement endpoints against one another and criteria related to assigning the likelihood of ecological impacts (e.g., none possible, likely, and probable) that the results from each measurement endpoint can be compared with on a stand alone basis. This process is not part of the EPA ecological risk assessment guidance. The perspective on the use of lines of evidence in the EPA ecological risk assessment guidance is discussed above.

Adherents to the establishment of *a priori* components in the WOE process claim such an approach does not deviate from the ERA process described by EPA but rather adds structure by defining and documenting the process. The Figure below shows the association of an *a priori* WOE approach process with the standard ERA process. Also shown in the column to the far right is the WDNR perspective on the use of a WOE approach in association with the ERA process. Such a WOE approach does not establish *a priori* weighting or criteria in the evaluation process. Uncertainties and advantages and limitations of measurements methods and endpoints used are normally discussed in the risk characterization section of the ERA.

<table>
<tr><th colspan="3">Association Between the Standard ERA Process and Steps Associated with Versions of the WOE Approach</th></tr>
<tr><th>EPA ERA Process</th><th>WOE Approach A Priori Established Components</th><th>WDNR Inclusive WOE Approach Consistent with EPA ERA Guidance</th></tr>
<tr><td>Problem Formulation
1) Develop a conceptual site model
2) Identify stressors (COPC)
3) Identify potential effects of stressors
4) Evaluate stressor release, migration, fate
5) Identify receptors
6) Identify exposure pathways
7) Select assessment endpoints (AEs)
8) Select measurement endpoints (MEs) based on relevance to AEs</td><td>1) Select MEs
2) Develop a numerical weight for each ME based on set of scaled attributes.
3) Develop criteria that relates the magnitude of difference in each ME between the study sites and reference site to the likelihood of degree of ecological impact (effect size ranges)
4) Recommendations have been made for a priori establishment of remedial actions/decisions based on ERA study findings, i.e., go beyond risk characterization and get into risk management.</td><td>1) Select MEs based on relationship to AEs and with known advantages and limitations based on literature and experience with use.
2) MEs selected consistent with the Problem Formulation process.</td></tr>
<tr><td>Exposure Assessment
1) Quantify release, migration, fate
2) Characterize receptors
3) Measure or estimate exposure points concentrations

Ecological Effects Assessment
1) Review of literature
2) Review of studies from sites with same COPC
3) Evaluate results from ME results (e.g., tox. testing, benthic studies)</td><td>1) Evaluate ME results considering the weight assigned to each ME and the magnitude criteria related to the assumed ecological impact for that ME as established a priori in the Problem Formulation stage.</td><td>1) Evaluate ME results along with pertinent literature and results from studies at other sites with same COPCs.
2) While magnitude of differences determined for each ME, effect sizes relatable to likelihood of impact based on individual ME is not done.
3) Exposure and effects assessment consistent with EPA ERA processes</td></tr>
<tr><td>Risk and Impact Characterization
1) Evaluate risks to AEs based on ME results
2) State and discuss observed and/or predicted adverse ecological effects
3) State and discuss Uncertainty and confidence analysis of methods, results, and information/data used in the risk characterization process and outcome.</td><td>1) Use the weight and magnitude criteria to evaluate risk for the site.</td><td>1) MEs and other info/data lines of evidence are used in a strength-of-evidence approach which is inclusive in nature and integrates and evaluates all pertinent information both ME results and other site and literature info and data. Logical evidence-based argument is for causation, potential impacts, and risk characterization. Uncertainty and confidence in methods and results is a part of the risk characterization discussion of the ERA.</td></tr>
<tr><td colspan="3">Comments on WOE Approach Using a priori Established Criteria and Components
Approach is mechanistic, inflexible, and exclusive of some data that would be useful in the characterization process. Prior knowledge of factors most relevant in population specific regulation is not available and cannot be developed a priori for generic application to all sites and situations. ERAs are a site-specific process based on integration of a number of sources of data and information subsequent to the collection, evaluation, and integration of that data. If a priori established criteria and components are to be used, it is antithetical to the ERA process that calls for thoughtful consideration and using casual arguments and inferences, logical arguments, professional judgements, and informed decision-making. If a priori criteria are established, why even go through the ERA process. Establishing a priori impact criteria based on stand alone ME results has no place in the ERA process for integrating and comparing all ME and other results for risk characterization decisions.</td></tr>
</table>

Suter's (1996) discussion of the WOE approach and the need for weighing the individual lines of evidence includes the following:

- If all the results of the lines of evidence are relatively consistent, then no formal weighting of the individual measurement endpoints is necessary.
- Recognition that attempts at quantification in the weighing of lines of evidence such as done by Menzie et al. (1996) may not give as reasonable a result in every case as a careful *ad hoc* weighing of all the data and information associated with the lines of evidence being used.
- Suter indicates that another approach to weighing multiple lines of evidence is to determine whether there are logical relationships among the lines of evidence. The process of developing a logical explanation for differences among lines of evidence is potentially more convincing than attempts at assigning quantitative weights to lines of evidence because of the mechanistic, inflexible nature of the process.
- Inferences of causality in ecological risk assessments are made by weight-of-evidence rather than traditional scientific standards of proof. The traditional standard for inference in science is, in effect, proof beyond a reasonable doubt in a decisive experiment. Such a standard is appropriate for pure science, which is engaged in adding to the body of reliable knowledge concerning the nature of the world. However, risk assessors do not have the luxury of suspending judgment until a scientific standard of confidence can be met. Decisions are made on schedules that are not within the control of scientists and will be made on other bases if scientific input is not available.
- Further research needs as noted by Suter include: 1) Proposed procedures for weighing evidence need to be validated; 2) Methods for presenting the results of weight-of-evidence analysis to risk managers and stakeholders need to be developed; and 3) Appropriate methods for estimating and expressing the uncertainties for different lines of evidence need to be developed.

Considering all of the above issues points to the need for assembling and interpreting (in a qualitative or quantitative fashion) all of the available data and information collected in the risk assessment process in an inductive process. The integrated data is then used in a logical evidence-based argument for causation and consideration of alternative hypotheses. Even within a formal experimental framework, a causal inference is made by logical assembly, presentation, and interpretation of available evidence including conclusions from statistical analyses. The basis for drawing conclusions from impact studies using causal arguments is more consistent with scientific method because it promotes thoughtful consideration about alternative hypotheses and about what constitutes a causal relationship. Thus causal inference by means of a carefully structured argument considering all pertinent data and information increases the likelihood of correct scientific conclusions. This approach is superior to one based on simplistic, mechanistic attempts to arbitrarily assign weighting factors to measurement endpoints in the initial formulation stages of the risk assessment and establishing a priori decision criteria that have not been validated and do not address the variables and interactions of factors occurring on a site-specific basis. Additionally such approaches are without provision or flexibility to examine all pertinent data and information that should be considered in the risk characterization process.

Some thoughts from Burton et al. (2002) on a WOE framework for assessing sediment contamination include:

- It is apparent that no single WOE approach is appropriate for all assessments of ecosystem impairment.
- Ideally, the various LOE will be collected in a synoptic, contemporaneous fashion to increase the certainty of exposure-effect linkages when the are integrated.
- An important aspect of the WOE process is to ensure that the data are valid prior to use and integration. Data validation should identify any questionable results and the questionable data identified with the appropriate qualifiers. By identifying questionable data, the uncertainties associated with the risk characterization outcomes can be minimized.
- In analyzing the LOE, aspects of the analyses that must be addressed and documented include: QA/QC; stressor magnitude; frequency; duration and interactions; and exposure-biological effect relationships. Estimation and evaluation of uncertainty of these aspects is a critical component of the analysis.
- Causality criteria used to link stressors and effects for each LOE should be clearly stated, together with how the links were established or refuted. The following key considerations for establishing strength of stressor effect linkages (causality) have been modified from USEPA (2000): co-occurrence (spatial correlation); temporality

(temporal correlation); effect magnitude (strength of link); consistency of association (multiple site; experimental confirmation (field or laboratory); plausibility (realistic stressor to response scenario); and specificity (stressor link to effect).

- Causality evaluation should be conducted both with individual LOE analyses, and following their integration, to establish how well each LOE links stressors with adverse biological responses. The approach uses results of quantitative assessments of stressor-effect relationships, followed by a combination of quantitative findings and best professional judgment in the causality decision. This is a diagnostic process whereby possible stressors are ruled out. While the causality criteria may be qualitative in nature, they add significant and reliable weight to the decision-making process when combined.
- Various processes have been described for establishing effect size limits (effect size means the amount of ecological change that is important enough to signal ecological concern), ranging from best professional judgment to more quantitative approaches. The effect limits establish at what level a measurement endpoint response is judged to be ecologically significant. The effect limits may be study specific, as they are affected by study design (e.g., characterization accuracy), societal values, and understanding of ecosystem components, dynamics, and inter-relationships. Effect size should be determined *a priori* even though it may be difficult to do so for biological effects.
- While it is possible to combine the information in multiple LOE into a single number that describes the degree of impairment can result in excessive reduction of information. Integration of all the lines of evidence into a single number is likely to over simplify the evidence.
- Each element of the proposed WOE framework comprises some degree of best professional judgement. Adequate expertise is required for the various LOE as well as quantitative methods used to evaluate the LOE. Expert judgement must be carefully incorporated and documented from the beginning to the end of the WOE process to ensure the transparency of the process.

Selecting and Weighing the Measurement Endpoint Attributes

EPAs ecological risk assessment guidance (1998) makes the following statement, part of which relates to the attributes of the individual measurement endpoints as lines of evidence that the risk assessor should consider in using the data involved in the overall WOE approach. This involves the degree and type of uncertainty associated with each measurement endpoint in its use as a line of evidence.

> *"There are three principal categories of factors for risk assessors to consider when evaluating lines of evidence: 1) Adequacy and quality of data, 2) degree and type of uncertainty associated with each line of evidence, and 3) relationship of the evidence to the risk assessment questions."*

Some schemes have been attempted (Menzie et al. 1996) to weight or establish the relative confidence level in each selected endpoint *a priori* in the problem formulation stage of risk assessments before the use of that endpoint in the overall WOE approach in the risk characterization stage of the risk assessment after all the data has been collected. This is part of the transparency of the risk characterization process that responsible parties and their consultants are asking for.

The workgroup involved in the Menzie et al. (1996) WOE approach identified three major components that reflect the weight-of-evidence of measurement endpoints, with respect to a specific assessment endpoint:

1) **Weight assigned to each measurement endpoint**. Measurement endpoints may vary in the degree to which they relate to the assessment endpoint, the quality of the data, or the manner in which they were applied. Based on these attributes, an investigator may assign more weight to, or have more confidence in one measurement endpoint compared to another.
2) **Magnitude of response in the measurement endpoint**. Strong or obvious responses are typically assigned greater weight than marginal or ambiguous responses.
3) **Concurrence among measurement endpoints**. More weight or confidence is generally attributed to findings in which there is agreement amount multiple measurement endpoints. An investigator generally has less confidence in findings in which the lines of evidence contradict one another.

In weighing the individual endpoints, Menzie et al. (1996) have a quantitative approach which assigns fixed numerical weights to ten attributes to reflect differing degrees of importance. The qualitative method does not involve preassigned weights but requires the risk assessor to rate endpoints in non-numerical terms (i.e., high, medium, and low). The qualitative approach is somewhat more flexible, in that it is more amenable to determining the relative importance of the attributes on a case-by-case basis. The risk assessor may opt either to assign weights on a case-by-case basis or to assume that each attribute is of equal importance. The rationale needs to be provided for the assigning the qualitative descriptor of weight or confidence in the use of each measurement endpoint in the qualitative method.

For the purposes of ecological risk assessments performed by and for WDNR, such approaches as the Menzie et al. (1996) quantitative approach will not be taken for assigning relative weights or confidence levels in the measurement endpoints to be reflective of the assessment. In cases were there is disagreement between risk assessors representative of the involved parties as to the importance of a particular measurement endpoint and the results from that endpoint, the attributes of Menzie et al. may be used and applied to the measurement endpoint in question to discern the basis of the perceived disagreements in measurement importance.

Weight-of-Evidence Decision Criteria Related to the Benthic Community Structure Line of Evidence

In interpreting the observed changes in benthic community structure, Batley et al. (2002) state the following:

> "In terms of the key processes in benthic sediments, can we identify that degree of change in community structure that is "unacceptable"? In other words, at what points should management action be triggered to protect the ecology of the system? Given that complete ecological understanding of these issues is lacking, a practical approach may be to consider any statistical detectable impact on community structure as ecologically undesirable."

Given the above, the results of the evaluation of the differences in metrics for benthic community structure in the following table are related to magnitude of differences in the metrics between the study sites and the reference site. Any statistically significant difference that is 20% or greater is deemed "ecologically undesireable". Given that the general lack of complete ecological understanding at this time about what the differences in individual metrics or multiple metrics may mean to the long term structure, function, and sustainability of the benthic community or to higher trophic level organisms that may consume benthic organisms who have bioaccumulated sediment contaminants, the > 20% difference level in any metric or multiple metrics will be considered of significance related to potential ecological impacts to the benthic community and will be integrated with other measurement endpoint results into the considerations in the risk characterization process..

No attempts will be made in the decision criteria in any stand alone line of evidence to attribute the degree of differences to the likelihood of increasing ecological or adverse impacts (e.g. the arbitrarily assignment of the qualitative descriptors such as none, potential, probable, and likely, to increasing ranges of differences). What will be assigned in the decision criteria for the benthic community metrics as shown in the following table is a qualitative descriptor related to the magnitude of difference in the metrics between the reference site and the study sites as follows:

% Difference in Metric Mean Between the Reference Site and the Study Site	Qualitative or Numerical Descriptor of the Magnitude of Difference Range
< 15%	Insignificant
> 15% - < 20%	Review
>20% - < 50	1
> 50% - < 75%	2
> 75% - < 100%	3
> 100%	4

The assignment of a descriptor for the likelihood of ecological impact to the benthic community assessment endpoint will only be done at that point in a matrix table where all lines of evidence are integrated, the magnitude of differences in all endpoints and metrics compared, and all pertinent information reviewed to make a structured argument for causation and characterizing risks.

Attempts to make final conclusions about the likelihood of ecological impacts as it involves each assessment endpoint through decision criteria, matrix tables, tabular matrixes, etc. are attempts to make understandable and relate a large amount of data. In doing so, the totality of information and data may not fit into such a system and some may not be used in consideration of its contribution to the overall weight-of-evidence. The approach we will take will not use the results of such matrix table displays alone to make final judgements as to the likely impacts to assessment endpoints. All of the pertinent data assembled for the site as it relates to measurement endpoints and other auxiliary and supportive data (e.g., other site and ERA studies and toxicological information) will be assembled and structured arguments made for causal inferences and the strength of the weight-of-evidence for the likelihood of impacts to the assessment endpoints.

Note the following table not only identifies statistically significant differences in metrics that are > 20% for consideration but also those differences that may or may not be statistically significant but are between 15% and also < 20%. All differences are assigned a qualitative descriptor of the magnitude of difference with those not related to statistical significance noted.

With the mean difference in benthic metrics in the following table, it is important to note the amount of difference is proceeded by either a positive or negative sign indicating an increase or decrease in the numerical value of the study site metric compared to the reference site metric. It is commonly associated that only decreases in such metrics such as abundance and diversity are associated with contaminant presence and ecological impacts. However, cases can be made that increases in certain benthic community metrics are also associated with contaminant presence and ecological impacts. Bias toward associating only reductions in metrics with impacts and overlooking the significance of increases would result in inaccurate estimations of impacts and risk in certain cases.

Weight-of-Evidence Criteria Related to the Toxicity Testing Line Of Evidence

The three different test species of macroinvertebrates utilized in the bioassays are representative of different life cycle characteristics, and habitats occupied in or on the sediment surface (infaunal and epi-benthic) by macroinvertebrates at the site. As such they have different a) life cycle lengths, b) reproductive potentials, c) exposure routes, d) death rates, e) sensitivities to the contaminants of concern, and e) different dispersal rates (i.e., immigration and emigration into and from a site). As with the benthic community metrics, there is an incomplete understanding of the ecological implications and/or consensus of what the endpoint results and effect sizes observed in the laboratory mean in attempts to extrapolate the results to the field setting. Therefore the same practical approach as applied to the benthic community metrics will be applied to the endpoint results from toxicity testing. This establishes that any statistically significant difference of 20% or greater between the survival and growth endpoints when the study site results are compared to the reference/control site results will be deemed ecologically undesirable. Differences of < 20% but greater than > 15% will also be evaluated, especially in those circumstances when the percentage difference is slightly less than 20% and statistically significant.

Given the above attributes of the different test organisms and variable responses, it will not be assumed that increasing ranges of differences in endpoint results between study sites and reference site have the same implications as related to potential ecological impact for all three test organisms. In just about all cases, prior knowledge of factors most relevant in population-specific regulation is not known. Attempts at establishing a priori effect sizes (i.e., amount of ecological change that is important enough to signal concern) generically to apply to all organisms and measurement endpoints (e.g., ranges of < 20%, > 20% to < 50%, and > 50%) is inappropriate, to mechanistic and inflexible, and may result in risk characterizations that are completely inaccurate.

As with benthic community metrics, no attempts will be made in the decision criteria in any stand alone line of evidence to attribute the degree of differences to the likelihood of increasing ecological or adverse impacts (e.g. the arbitrarily assignment of the qualitative descriptors such as none, potential, probable, and likely, to increasing ranges of differences). What will be assigned in the decision criteria for the toxicity testing endpoints as shown in the following table is a qualitative descriptor related to the magnitude of difference in the toxicity test endpoints between the reference site and the study sites as follows:

% Difference in Metric Mean Between the Reference Site and the Study Site	Qualitative or Numerical Descriptor of the Magnitude of Difference Range
< 15%	Insignificant
> 15% - < 20%	Review
>20% - < 50	1
> 50% - < 75%	2
> 75% - < 100%	3
> 100%	4

The assignment of a descriptor for the likelihood of ecological impact to the toxicity testing endpoints will only be done at that point in a matrix table where all lines of evidence are integrated, the magnitude of differences in all endpoints and metrics compared, and all pertinent information reviewed to make a structured argument for causation and characterizing risks.

Attempts to make final conclusions about the likelihood of ecological impacts as it involves each assessment endpoint through decision criteria, matrix tables, tabular matrixes, etc. are attempts to make understandable and relate a large amount of data. In doing so, the totality of information and data may not fit into such a system and some may not be used in consideration of its contribution to the overall weight-of-evidence. The approach we will take will not use the results of such matrix table displays alone to make final judgements as to the likely impacts to assessment endpoints. All of the pertinent data assembled for the site as it relates to measurement endpoints and other auxiliary and supportive data (e.g., other site and ERA studies and toxicological information) will be assembled and structured arguments made for causal inferences and the strength of the weight-of-evidence for the likelihood of impacts to the assessment endpoints.

Establishing *a priori* decision criteria based on stand alone toxicity testing results are presumptuous unless comparable types and amounts of laboratory toxicity endpoint impacts have been related to site ecological impacts involving benthic community assessment endpoints. The demonstration of correlations between laboratory-based toxicity test results and field impacts to the benthic community assessment endpoint must from other studies and from other sites. These correlations at other sites may be at best tenuous in their application at other sites because of the myriad of site-specific variables involved that will affect results. Therefore, other than determining the magnitude of differences between metrics and measurements at this stage, we believe it is inappropriate to try associate stand alone metrics and measurements to the likelihood of ecological impacts at this point.

Weight-of-Evidence Decision Criteria Related to the Sediment Chemistry and Sediment Quality Guideline Lines of Evidence

The table below summarizes the sediment chemistry results and sediment quality guidelines for a number of measurements related to the mix of petroleum hydrocarbons that are the primary contaminants of concern for the site. The sediment chemistry measurements (as opposed to the sediment quality guideline values) are not risk based but simply represent the relative degree of enrichment of the measured chemicals in the study site sediments compared to the reference site. To be consistent with how percent differences in the benthic community metrics and toxicity testing endpoint results were calculated, the percent differences in the chemical concentrations and guideline values between the study sites and the reference site were calculated similarly yielding relative percent difference values.

% Difference in Metric Mean Between the Reference Site and the Study Site	Qualitative or Numerical Descriptor of the Magnitude of Difference Range
< 15%	Insignificant
> 15% - < 20%	Review
>20% - < 50	1 - Low
> 50% - < 75%	2 - High
> 75% - < 100%	3 - High
> 100%	4 -Very High

The percent difference value as calculated applies to concentration values as well as some other metrics (e.g. PEC-Q, ESG and number of TEC and PECs exceeded). All differences are related to magnitude of differences in chemical concentrations and other metrics and are not related to impacts to the benthic community assessment endpoint. Decisions and decision criteria about the significance of the chemical test results can only be made in the total context of integrating all the measurement endpoint data and other auxiliary and supporting information. Establishing *a priori* decision criteria based on stand alone toxicity testing results are presumptuous unless comparable types and amounts of laboratory toxicity endpoint impacts have been related to site ecological impacts involving benthic community assessment endpoints.

ATTACHMENT 3

Draft

Remediation and Redevelopment Program Review of the Basis For Establishing the Risk-Based Preliminary Remediation Goals For the Coal Tar-Contaminated Wood and Sand Substrates In the Lakefront Embayment, Ashland, Wisconsin

Summary

The U.S. EPA CSTAG met in Ashland July 15-17, 2002 to review the Ashland MGP investigation results to ensure that their eleven risk management principles for managing contaminated sediment risks at hazard waste sites were followed. The CSTAG will produce a series of recommendations related to each of the eleven risk principles. Based on the CSTAG discussions of July 17, they had some preliminary recommendations as it related to the ecological risk assessment (ERA) performed for the contaminated sediments and surface waters in the Lakefront embayment and specifically for the use of the bioassay lines of evidence in the derivation of the preliminary remediation goals (PRGs). In anticipation of the CSTAG recommendations in regard to the ERA, the bioassay results, and the PRGs, a review was undertaken of the use of the bioassay results obtained from two years of testing the site media.

Based on the results of the bioassay testing and all derived metrics from this data as contained in the ERA and Supplemental ERA for the Ashland site, an attempt was made to convert the metrics into a more quantitative expression of risks based on protectiveness levels to organisms and endpoints.
In this fashion, risk managers would be better able to relate the lower and upper effect bound thresholds associated with the preliminary remediation goals to incremental increases in the total PAH coal tar contaminant concentrations. All effect-based concentrations and protectiveness levels are based on organic carbon normalized (NOC) TPAH concentrations. The NOC TPAH concentrations are converted to dry weight concentrations based on a representative range of total organic carbon concentrations (TOC) found in the reference substrates at the site.

An array of metrics derived from the bioassay results were looked at in order to derive the protectiveness levels for the organisms and endpoints. The metrics included NOAECs, MATCs, LOAECs, 10-d LC50 concentrations, LC10 concentrations, and chronic toxicity values derived from the LC50 concentrations. The protectiveness levels derived from the metrics are interpolated to apply to the field organisms and populations based on the need to protect the assessment endpoints chosen in the ERA. Protection of the assessment endpoints ensures that the stability and functioning of the community of aquatic organisms within the Lakefront embayment.

Along with the metrics derived from the results of the bioassays and the establishment of levels of desired protectiveness for organisms and endpoints related to the assessment endpoints, the metrics were also compared with two toxicity benchmarks to determine the predictability of the toxicity benchmarks based on the site-specific toxicity testing.

The organic carbon normalized TPAH concentrations related to protectiveness levels of organisms and endpoints that should be used as the lower effect bound threshold concentration for the preliminary remediation goal is 100 –200 ug TPAH / g OC which translates into dry weight concentrations of 2 to 20 mg TPAH / kg of wood or substrate based on a range of organic carbon in the substrates of from 2 to 12 %.

Introduction

The U.S. EPA Contaminated Sediment Technical Assistance Group (CSTAG) met in Ashland on July 15-17, 2002 to review the Ashland Manufactured Gas Plant (MGP) investigation results for the operational units of the site. The CSTAG members are Superfund project managers from around the United States. The CSTAG was formed within EPA to review Superfund sites to assure consistency where contaminated sediments are involved. The CSTAG review focused on the compatibility of the work performed to date at the site with U.S. EPA guidance (2002) based on 11 principles for managing contaminated sediment risks at hazardous waste sites. The recommendations that the CSTAG will produce from their review will be based on recommendations to ensure the 11 risk management principles are met by the all work products from the site from the RI through the FS and remedy selection.

Based on the CSTAG discussions on July 17, there was a preliminary indication that they would be developing a number of recommendations in association with two of the risk management principles associated with the ecological risk assessment performed for the contaminated sediments. The two risk management principles involved were numbers, 6 and 8 that were:

6) Carefully evaluate the assumptions and uncertainties associated with the site characterization data and site models.

8) Ensure that sediment cleanup levels are clearly tied to risk management goals.

<u>CSTAG Discussions and Possible Recommendations In Regard to Principles 6 and 8</u>

Overall, it is believed by this reviewer that many of the points raised in the CSTAG discussion that may find their way into their recommendations could be addressed by a complete read of the existing 1998 ecological risk and 2002 supplementary documents and series of response made to commentators on the documents.

1) The measurement and assessment endpoints, site model, and uncertainties associated with the lines of evidence used are identified in the ERA documents.
2) The extent that the fish were used as an assessment endpoint focused on impacts to larval fish as demonstrated by exposures to contaminated site media.
3) The possible recommendation for additional literature reviews and literature reviews of current research of toxicity in the Great Lakes is superfluous in light of literature reviews already done for the ERA. In the absence of the CSTAG members not having read the complete ERA-related documents, it is not seen how they would make this or some of their other potential recommendations. The recommendations possibly will lead to redundancy due to the issues already been dealt with in the ERA documents.
4) Given the literature reviewed for the ERA, it would also not seem that a potential CSTAG recommendation for contacting the ORD-Nargansett Lab for information on toxic effects to fish is needed. The ERA carried the fish-related assessment endpoints to the point wanted and additional studies in this area are not deemed necessary. After a read of the ERA, the CSTAG needs to identify specific needs and additional studies related to site fish. Any methodologies need to be able to demonstrate more than simple exposure
5) While it would be interesting to contact the RPMs of other PAH sites to determine what the cleanup goals are for their sites, along with the numerical cleanup goals they may provide has to be a complete explanation and rationale for developing and choosing those numbers. My experience with Superfund sites in the past is that there is a disconnect between scientifically-based numbers derived through risk

assessments and other means and the cleanup numbers that end up in RODs and Consent Decrees. They need to demonstrate to me that all the cleanup numbers are scientifically based, fully explained, and arrived at through consistent processes.

6) As to consulting with additional technical support in the selection of cleanup goals, it was not specified by CSTAG who constitutes this additional technical support. The technical support used to date has been reference to current publications and contacts within USGS and EPA that are involved in associating TPAH mixtures with effects to aquatic organisms and the UV enhancement of PAH toxicity. The CSTAG group discussed and raised a number of issues in association with the role that the bioassay lines of evidence played in the establishment of the preliminary remediation goals (PRGs) of 2 –20 mg/kg dry weight TPAHs. To respond to the issues raised in regard to the use, explanation, and uncertainties associated with the bioassays as lines of evidence in establishing the PRGs, the bioassay results in the ERA were re-examined for the information needs as stated by CSTAG. As necessary further explanations and rationales were developed and are provided in the sections that follow. All of the information and data used below is contained in the 1998 ERA and 2002 supplement.

Risk Description Basis For Deriving the Preliminary Remediation Goals For TPAH Concentrations in the Substrates Present Off Kreher Park

Selection of Assessment Endpoints as Established in The ERA (SEH, 1998 and 2002) For the Ashland Lakefront Property Contaminated Sediments

Based on identified potential ecological receptors and the result of the exposure pathways evaluation, the following assessment endpoints were identified during problem formulation dealing with the health of the following aquatic organism communities:

- Survival and growth of benthic organisms
- Benthic community health
- Survival and growth of fish and water column organisms (phytoplankton and zooplankton)

The health and vitality of the infaunal, epibenthic, and water column organisms are of ecological value because the organisms associated with the trophic levels involved are important for the structure and functioning of the localized and wider area near shore aquatic ecosystems.

Risk Description (from U.S. EPA, 1997)

A key to risk description for Superfund sites is documentation of environmental contamination levels that bound the threshold for adverse effects on the assessment endpoints. The risk description identifies the threshold for effects on the assessment endpoints as a range between contamination levels posing no ecological risks and the lowest contamination levels identified as likely to produce adverse ecological effects

Threshold for Effects on Assessment Endpoints

Key outputs of the risk characterization step are contaminant concentrations in the substrates that bound the threshold for estimated adverse ecological effects given the uncertainty inherent in the data and models used. The effect threshold bounds are described as:

a) The lower bound of the threshold would be based on consistent conservative assumptions and NOAEC toxicity values.
b) The upper bound would be based on observed impacts or predictions that ecological impacts could occur. This upper bound would be developed using consistent assumptions site-specific data, LOAEC toxicity values, or an impact evaluation.

Considerations Made For Deriving PRGs From 1998 and 2001 Sediment Bioassay Data

The preliminary remediation goals (PRGs) based on the bioassay lines of evidence are meant to characterize the threshold bounds for protecting the assessment endpoints. The 1998 and 2001 bioassays include test organisms representative of various a) trophic levels, b) habitats, c) reproductive potentials, d) life cycle lengths, e) sensitivities to certain bioaccumulated PAHs by UV light, and f) sensitivities to PAHs in general. Bioassay results are available from two different years and involve a number of exposure methodologies i.e., exposures to whole undiluted sediments in 1998 and exposure to sediments based on a serial dilution methodology in 2001; exposure of test organisms to water over contaminated sediments; exposures of test organisms to sediment elutriates and serial dilutions of sediment elutriate; and extraction of benthic organisms from sediment and exposure to UV light after normal test lengths under lab lights. The testing and exposure methodologies were applied to the two primary contaminated substrates at the site that include the natural sands and silty sands and the mobile wood waste-associated substrate that overlays the sands. Similar substrates located lakeward just beyond the boundary of the coal tar contaminated area were selected for use as reference sites in the testing.

A combination of five different test organisms were used during the toxicity testing. The organisms, test lengths, endpoints, and media tested involved are summarized below. The testing protocols for all the methodologies are contained in the initial and supplemental ERAs for the site.
The test methods and results for both years met all testing protocols and test acceptability criteria and all tests and results for both years were given equal weight when combined to derive effect levels.
In the Test Organism column below, an abbreviated descriptor for the test that is used in the discussion below is shown. The 1998 testing involving *Pimephales promelas* in serial dilutions of substrate elutriate is discussed in a separate section below and is not combined with the results from all the other tests in determining ranges of protectiveness levels. However the results have equal pertinence for consideration of the threshold effect bounds.

Year	Test Organism	Test Duration	Test Endpoints	Media Tested
1998	*Hyalella azteca* HA-10	10-d	Growth Survival	Sand and Wood Substrates
2001	*Hyalella azteca* HA-28	28-d	Growth Survival	Sand and Wood Substrates
2001	*Hyalella azteca* HA-28 UV	28-d - UV Light	Survival Growth	Sand and Wood Substrates
1998 and 2001	*Chironomus tentans* CT-10	10-d	Survival Growth	Sand and Wood Substrates
1998	*Lumbriculus variegatus* LV-10	10-d	Survival Growth	Sand and Wood Substrates
1998	*Lumbriculus variegatus* LV-10 UV	2 hrs UV light after 13-d	Survival	Sand and Wood Substrates
1998	*Pimephales promelas* (fry)	7-d	Survival Growth	Substrate Elutriates Serial Dilutions
1998	*Pimephales promelas* *(fry)*	4 hrs UV Light after 24-h	Survival	Substrate Elutriates
2001	*Pimephales promelas* PP-7 (fry)	7-d	Survival Growth	Water Over Substrates
1998	*Daphnia magna* DM-48	48 hrs	Survival	Substrate Elutriates

1998	*Daphnia magna* DM-48 UV	4 hrs UV Light after 48 hrs	Survival	Substrate Elutriates

Importance of the Use of Organic Carbon Normalized Concentrations of Total PAHs (NOC TPAHs ug/g OC)

The common element used to relate all the endpoint results and effect levels in the following discussion is the conversion and expression of all TPAH concentrations involved for each site tested on an organic carbon normalized basis. The organic carbon content of sediments is the primary determiner and controller of the mobility and bioavailability of PAHs to aquatic organisms.and better serves to facilitate site-to-site comparisons of TPAH concentrations and effects (Di Torro, D.M. et al. 1991). Expression of TPAH concentrations on an organic carbon normalized basis also allows comparisons with NOC TPAH concentrations and related effects from other studies. The U.S. EPA's ESG Technical Basis Document demonstrates that biological responses of benthic organisms for nonionic organic chemicals such as PAHs in sediment are different across sediments when the sediment concentrations are expressed on a dry weight basis but similar when expressed on a
ug chemical/g organic carbon basis. In each section below where effect levels and percentage of protectiveness to organism/endpoints are discussed, the NOC TPAH concentrations are converted to dry weight (mg/kg) concentrations based on a representative range of total organic carbon concentrations found in the reference site sediments.

In the discussion below, the results from both years of toxicity testing are combined, with the effect-related NOC TPAH concentrations from the tests ordered from lowest to highest. Each of the ordered concentrations are related to the percent of the organism/endpoints that would be protected if all the substrates with concentrations greater than the specific ordered concentration were removed. This in effect provides threshold effect bounds that would help risk managers relate various levels of protection to the benthic and water column assessment endpoints established in the ERA. No attempts were made to discern why some tests show different sensitivities between years for some species and between the different substrates. When all results are combined and the effect-related concentrations established, it is believed that the test organisms, that are serving as surrogates or are representative of taxa at the site, show the range of sensitivities, life cycles, trophic levels and habitats of the aquatic organisms at the site under the various exposure scenarios.

The relevancy of exposure of organisms to the UV component of sunlight and sediment elutriates are discussed in the ERA and in the past responses to comments on these aspects and will not be repeated here. The results of tests involving the UV light test component are included into the combined testing results for both years as identified with the UV descriptor (e.g., HA-28 UV). It is important that the non-UV and UV light test results be reviewed and compared in order to see the implications associated with the photoacitivated toxicity of certain PAH compounds that have been bioaccumulated within the tissues of aquatic organisms. For example, based on the 2001 *Hyalella azteca* bioassay results where LC50s were calculated (TPAH NOC basis), exposure to UV light increased toxicity approximately 3.7 times based on exposures to the contaminated sands and 3 times based on exposures to the contaminated wood substrates. In 1998, *Lumbriculus variegatus* and *Daphnia magna* experienced 100% mortality after being exposed to the contaminated media for a period of time and then exposed to UV light

Use of the Serial Dilution Methodology For Sediments In the 2001 Bioassays

In 2001, wood and sand substrates were collected from the most contaminated areas of the site along with clean reference sediments. The reference site sediments were mixed in proportions to create the following serial dilutions for both substrates: 100 % contaminated substrate, 50%, 25%, 10%, and 1%. The serial dilutions created a range of incremental concentrations of the coal tar contaminants in the substrates for use in the bioassay exposures. As identified in the uncertainty analysis of the supplemental ERA, resulting chemical concentrations in the serial dilutions where not in proportion to the percent of dilution. On a relative basis, the test organisms did not appear as sensitive in the 2001 bioassays compared to the 1998 bioassays that were ran on undiluted, bulk sediments collected from the site. There are some uncertainties associated with the serial dilution methodology as it applies to sediments. Some of these uncertainties as identified by Burton (1991), Nelson et. al (1993), and Giesy et al. (1990) include:

- Currently, little information is available o the most appropriate method for diluting test sediments to obtain a graded contaminant concentration or concerning the methodological effects of such dilution.

- Adding clean sediments to contaminated sediments increases fresh sorptive sites for contaminants (e.g., TOC in clean sediments acting as a sorbent), thus reducing the biological effect further than simple dilution.
- In all dilution methods, both the effect of contact time of the interstitial water and sediment (i.e., equilibrium) and the effects of disrupting the sediments integrity on toxicant availability must be considered.
- The diluent reference substrate may alter sediment porosity, thus altering availability and desorption/adsorption kinetics.

While the sediment dilutions may reduce the contaminants present in a proportionate manner, they may change the kinetics of the contaminant behavior to make it less mobile and less available than expected. This would result in an underestimate of actual exposure risks and make the organisms appear less sensitive or more tolerant of higher concentrations of contaminant in the sediments. If this is what occurred in the 2001 bioassays using the serial dilution methodology, the sediment concentrations associated with effects in the laboratory may be an underestimate of the actual exposure risks to organisms exposed to comparable bulk sediment concentrations in the field.

Relationship of the Bioassay Endpoints Results Used Develop the Preliminary Remediation Goals to Ecological Significance

As discussed above, an array of test organisms of various characteristics were exposed to different contaminated site media and reference site media in the two years of testing. The test organisms of a range of sensitivity to the site contaminants and exposures in the sampled site media are judged to be representative of the organisms in the field populations and their sensitivities. The stated point in the ERA that endpoint effects were associated with the possibility of adverse ecological effects was any statistically significant reduction of 20% or more in growth, biomass, or survival when the study site results were compared to the reference site results. The 20% value in bioassay results related to ecological effects in the field is based on Lawrence (1999), Michelsen, 1999), Suter (1996), Suter and Tsao (1996), and Chapman et al. 1997).

The 20% value for differences in bioassay endpoints between the study sites and the reference sites is the best reasonably conservative estimate that, when extrapolated to the field, would result in protection of populations to maintain the stability and function of the localized ecosystem involved in the near shore area. This attempts to consider and in light of not knowing all the organism interrelationships involving variable life cycle lengths, reproductive capacities, differential tolerances for natural and contaminant stressors, genotypic composition of the populations, and locations in the food web where redundancy and immigration, or both, are low.

The population dynamics and numbers of aquatic organisms within the near shore area off the Lakefront property will be controlled by a number of intrinsic (factors within populations) and extrinsic factors (primarily annual and seasonal differences in the physical environment or nonbiological factors). To date, ecologist have not been able to distinguish quantitatively between intrinsic and natural extrinsic factors that may cause various types of population density variations (Odum,1966). Overlain on these natural factors in affecting the population dynamics is the presence of coal tar contaminants and oils in the site media. The presence of the contaminant stressors could confound the predictability of population control factors and resulting population levels. It has only been recently that research has looked at the relationships between toxicants as environmental stressors and population dynamics (Barnthouse, 1993). Qualitatively, the response of a population to a toxicant contaminant stressor is influenced by the preexisting pattern of natural environmental variability, the age-specific survival and reproduction of organisms, and the intensity and duration of contaminant exposure. Quantitative conclusions about the combined effects of both natural and contaminant stressors require case-by-case studies of populations at the site of interest.

Role of Background or Reference Site Concentrations of TPAHs and Total Organic Carbon

The points of reference both for the results of the toxicity testing and the TPAH and organic carbon contents of the study site substrates were the reference site wood and sand substrates. Both sites were selected lakeward beyond the identified boundary of the coal-contaminated substrates in the Lakefront embayment. Subsequent toxicity testing in 1998 showed effects to some endpoints in the sand and wood substrates designated as the reference sites. It was surmised that due to the locations that were separate but near to the coal tar contaminated boundary, that movement of contaminants had occurred between the contaminated area and more lakeward areas over time. Given the wind and wave activity in the high-energy near-shore area involved, it can be expected that mobilization of coal tar contaminants in the substrates will occur with subsequent transport on and off site. The basic characteristics of the reference site substrates are shown below.

Characteristics of the Reference Site Substrates Used In the 1998 and 2001 Bioassay Testing						
Year	**Sand**			**Wood**		
	Dry Wt. TPAH ug/kg	**% TOC**	**NOC TPAH ug/g OC**	**Dry Wt. TPAH ug/kg**	**% TOC**	**NOC TPAH ug/g OC**
1998	424	0.46 %	92	6,543	5.7 %	584
2001	3,110	2.1 %	148	12,597	19 %	66
Average	1,767	1.28 %	120	9,570	12.35 %	325

The organic carbon content of the reference sites is assumed to originate from natural materials, i.e., the wood reference site would have degraded wood material acting as the sorbent and the sand would have organic materials from such sources as phytoplankton that settled out of the water column. Given the nearness of the sand reference site to the wood area, organic carbon from the wood area may be present at the sand reference site. The study sites would have the same sources of natural organic carbon. The additional source of organic carbon at the contaminated wood and sand reference sites is the coal tar oils. However, the coal tar oils are not the innocuous source of organic carbon as are natural organic carbons. The coal tar oils in the form of NAPLs or DNAPLs associated with the sand and wood substrates are acting as both a sorbent and a source of PAHs to the organisms. It needs to be followed up on if the equilibrium partitioning model based on natural organic carbon overestimates or underestimates the PAHs partitioning from coal tar organic carbon. If the model used underestimates the PAHs that are being partitioned from coal tars, then the possible effects on the organisms and endpoints may also be underestimated as there maybe a greater aqueous concentration of PAHs present than predicted from the model.

Establishment of NOAEC, MATC, and LOAEC Concentrations and Their Use As a Line of Evidence In Deriving the Preliminary Remediation Goals Related to the Lower and Upper Bound Effect Thresholds Established in the ERA

The data generated in toxicity tests where a test organism is exposed to increasing range of contaminant concentrations in samples of the site media (pore water, sediments, water over sediments, and elutriates of sediments) enables the determination of a *no observed adverse effect concentrations* (NOAEC), which is the highest test concentration of contaminant evaluated in the media tested that causes no statistically significant difference in effect in exposed organisms compared to the mean response at the control site media or reference site media in a specific test.

The *lowest observed adverse effect concentration* (LOAEC) is the lowest test concentration evaluated in the site media that has a statistically significant adverse effect on the exposed organisms compared to the mean response at the control site test media or the reference site media in a specific test. In the Ashland Lakefront ERA, a statistically significant result that represented a 20% more reduction in the growth, survival, or reproductive endpoints in the study site sediments compared to the reference site sediment results were interpolated to have ecological significance to populations in the field as discussed elsewhere in these comments.

A *maximum acceptable toxicant concentration* (MATC) can be estimated. This is an estimated threshold concentration of a chemical within a range defined by the highest concentration tested at which no significant adverse effect is observed (NOAEC) and the lowest concentration tested at which some significant deleterious effect was observed (LOAEC). Since it is not possible to test an unlimited number of intermediate concentrations, an MATC is generally reported as being greater than the NOAEC and less than the LOAEC. The MATC value has been derived in some studies by deriving the geometric mean of the NOAEC and LOAEC values (U.S. EPA, 1998).

Uncertainties Associated With Using NOAEC, LOAEC, and the MATC Values For Derivation of Effects Thresholds

▪ The perceived advantage of the NOAEC relative to regression-derived estimates, such as the LC50, is that is it is easy to calculate, easy to understand, and is an important summary statistic in current ecotoxicity testing and chemical risk assessment procedures. Its derivation does not require the assumptions of a specific model. However, its application does carry the assumption of a toxic threshold: that there is no effect below some threshold concentration. This threshold concentration is often assumed to be between the NOAEC and the LOAEC.

▪ The establishment of the NOAEC and LOAEC values will depend on the number of different contaminant concentrations tested. If there are large range of concentrations between two concentrations used in the test, there will be uncertainty as to where the actual NOAEC or LOAEC value is located. To get closer to the actual NOAEC and LOAEC concentrations, a series of range finding toxicity tests would need to be conducted, with each test round narrowing the untested ranges between two original test concentrations. With limited test point concentrations, the actual NOAEC value may actually be less than the value found. By the same token, the LOAEC could also be less if concentrations between two tested points were tested. The NOAEC and LOAEC values are artificial constructs and should not be used on a stand alone basis in establishing protective values but should be used along with other data and other effect-derivation processes.

▪ As reviewed by Crane and Newman (2000) there are disadvantages and uncertainties with the use of NOAEC values to establish contaminant concentration thresholds protective of aquatic organisms.
The major conclusion of the Crane and Newman (2000) analysis suggests that the NOAEC value is neither a consistent summary statistic nor an indicator of safe concentrations of toxic chemicals in all situations. Summarizing ecotoxicity data as a NOAEC can provide a compromised picture of chemical safety in the environment. In most cases, a risk assessor using NOAEC values will have no way of knowing whether these values are indicative of low, medium, or high effects on the endpoint of interest, but the NOAEC is rarely if ever is an indicator of no effect.

▪ Other disadvantages identified by Chapman et al. (1996) for the derivation and use of NOAEC as a summary statistic include:

a) The NOAEC must be one of the concentrations used in an experiment since hypothesis testing does not allow interpolation between test concentrations. Thus, an important determinant of the NOEC is the choice of test concentrations.
b) The NOAEC tends to increase as the precision of the experiment decreases. A poorly performed

experiment may yield a large NOAEC that implies a safer chemical when in fact it is not safe.

c) Confidence intervals cannot be calculated for NOAECs. It is therefore not possible to compare the accuracy of NOAEC values from different experiments.

d) NOAECs may occur at concentrations which actually cause large effects because high experimental variability reduces statistical sensitivity, thus preventing these effects being detected as statistically significant. The NOAEC therefore cannot be considered an estimate of a safe dose.

Similarly, as noted by Newman (1995) cited in Newman et al. (2000), the LC50, NOAEC, and maximum acceptable toxicant concentration (MATC) have very significant deficiencies as measures of effect to field populations and communities. Any secondary metric based on such compromised metrics possesses the same deficiencies.

Using the NOAECs and LOAECs to Derive NOC TPAH Concentrations Protective of Variable Numbers of Bioassay Organisms and Endpoints

The uncertainties and critiques of the derived NOAECs and metrics associated with them as identified above were kept in mind in the derivation of the PRGs for the Lakefront sediments. A process was employed that looked at and compared using the NOAEC values and secondary metrics associated with it. The two secondary metrics included deriving concentrations protective of organism/endpoints based on 1) taking a mean value between the NOAEC value and the next most lowest test concentration for the substrate tested to derive an adjusted NOAEC protective value (ANPV), and 2) deriving a MATC concentration based on taking the mean of the NOAEC value and the LOAEC for the substrate tested.

Derivation of Maximum Acceptable Toxicant Concentrations (MATC) For TPAHs and Relationship to Organism/Endpoint Protectiveness

MATC values were calculated by taking the geometric mean of each the 1998 and 2001 toxicity tests for which both a NOAEC and LOAEC value could be derived (15 of 18 toxicity tests). Table 1 shows the MATC TPAH NOC ordered concentrations related to each organism/endpoints. To protect 80 % or more of the organism/endpoints, the TPAH concentration has to be 236 ug/g OC or less. Another way of looking at the protectiveness percentages in the far right column is to assume all concentrations of TPAHs in the substrate greater than the concentration associated with the desired level of protectiveness have to be removed.

Table 2 converts the TPAH NOC concentrations into dry weight concentrations based on a representative range of TOC concentrations at the site as represented by the reference site content.

<u>Use of the NOAEC Values For TPAH Concentrations and Relationships to Organism/Endpoint Protectiveness</u>

Table 3 shows the derived NOAEC and LOAECs for the four 2001 toxicity tests ran on both the contaminated sand and wood sites and their reference sites.

Table 4 shows the lowest NOAEC and LOAECs when the values from both substrates are combined from the 2001 toxicity tests.

Table 5 shows the lowest NOAEC and LOAECs when the results from the five 1998 toxicity test ran on the sand and wood substrates are combined and ordered by increasing TPAH NOC concentrations.

Table 6 shows the combined lowest NOAEC and LOAECs values from both the 1998 and 2001 toxicity tests ordered by increasing TPAH NOC concentrations .

Table 7 shows the secondary metric designated the Adjusted NOAEC Protective Value (ANPV) which was derived by taking the mean of the lowest NOAEC values and the next most lowest TPAH concentration tested. This was done based on the above discussion that the NOAEC value may in fact not be a protective value. The derived ANPVs are shown on Tables 3,4, and 5 in relationship to the NOAEC values. The ANPV values in Table 7 are ordered by increasing TPAH NOC concentrations.

Tables 6 and 7 show in the column to the right the decreasing level of protectiveness to the afforded the organisms/endpoints as the TPAH NOC concentrations increase. As discussed above, it is assumed all substrates with TPAH concentrations greater than the concentration associated with the desired protectiveness level are removed.

Tables 8 (all NOAEC), 9 (Lowest NOAEC), and 10 (Adjusted NOAEC) convert the TPAH NOCs to dry weight concentrations based on a representative range of total organic carbon that may be found associated with the reference sites.

Table 8 is based on combining the NOAECs from all of the 1998 and 2001 toxicity testing without regard to the lowest values between years or substrates tested.

Table 9 is based on combining the lowest NOAECs from the combined 1998 and 2001 toxicity testing.

Table 10 is based on the adjusted NOAEC values (ANPV) derived from the lowest NOAEC from the combined 1998 and 2001 toxicity testing.

Use of LC_{50} Concentrations As a Line of Evidence In Deriving Preliminary Remediation Goals Related to the Lower and Upper Bound Effect Thresholds Established in the ERA

The Median Lethal Concentration (LC_{50}) is a statistically or graphically estimated concentration that is to be expected to be lethal to 50 % of a group of organism under specified conditions. LSRI (2001) used U.S. EPA (2000) guidance to determine LC50 concentrations from the results of the 2001 bioassays ran on the sand and wood substrates. The 10-d LC50 values determined by LSRI for the bioassays using the three test species on each substrate are shown in Table 12.

The large effect sizes associated with LC50 values limit their utility in risk assessments unless for use in establishment of an upper effect bound threshold for the prediction with certainty that there will be adverse effects on the survival assessment endpoint and therefore ecological impacts will occur if the LC50 values are approached or exceeded. LC50 values should not be used in the consideration of remediation goals for a site. Short term acute toxicity tests as represented by the LC50 value are not sufficiently sensitive to detect the early stages of ecosystem stress. Significant effects on populations can occur at much lower concentrations of contaminants than LC50 values. Longer term studies have shown how species populations such as amphipods can suffer eventual extinction at contaminant levels below those that effect survival as measured by LC50 values (Ingersoll et al. 1997). Short term acute toxicity tests as exemplified by LC50 values are not sufficiently sensitive to measure or detect the early stages of stresses to populations that can occur from chronic, longer-term exposures to lower contaminant concentrations.

Two effect-related values can be derived from the LC50 values that may have use in determining lower chronic effect-related concentrations or a lower level of lethality more related to sustaining populations of organisms.

Application Factor Applied to LC50 Values To Derive a Chronic Toxicity Value

A value characterizing chronic toxicity can be derived from an LC50 value using an adjustment or application factor. Based on U.S. EPA (1985), "the acute-chronic ratio expresses the relationship between the concentration of an effluent or a toxicant causing acute toxicity to a species and the concentration off a toxicant causing chronic toxicity to that same species. The acute-chronic ratio has commonly been used to extrapolate to a "chronic toxicity" concentration using an available acute toxicity data point. The most widely used acute-chronic ratios are 20 and 100, i.e., the chronic toxicity concentration is 1/20 to 1/100 of the acute toxicity level. Commonly 20 has been used for non-persistent toxicants and 100 has been used for persistent toxicants. These numbers have been used for both chemical-specific and whole-effluent approaches to impact assessment. The acute-chronic ratio is a source of variability in assessing toxic water quality impact because the ratio varies both between species and, for any one species, between different toxicants." Depending on the species, the adjustment factor used in various ERA's ranged from 10 to 300 (Duke et al. 2000). U.S. EPA recommends regulatory agencies use a 10:1 ratio or the acute to chronic ratio in the absence of specific information for whole effluents. U.S. EPA (Draft, 2000) uses an acute to chronic ratio of 4.16 for TPAHs as it relates to sediments.

The LC50 values in Table 12 below were adjusted using an assumed acute:chronic ratio of 4.16:1. The 4.16 ratio translates into the chronic value being 24% of the acute LC50 value. 24 % LC50 concentrations for each test organism on each substrate are also shown in Table 12. Table 13 shows the NOC TPAH concentrations for the 24 % LC50 values ordered in increasing concentrations as they relate to each test result. The column to the right in Table 13 shows what percentage of the total organism/endpoints would be protected assuming a cleanup of sediments occurred to that concentration and all concentration greater than that concentration were removed.

LC10 Value

Based on the LSRI LC50 value and the survival data for each test and substrate in the LSRI report, estimated LC10 values were graphically derived and are shown below in Table 14. Table 15 shows the NOC TPAH concentrations for the LC10 values ordered in increasing concentrations as they relate to each test result. The column to the right in Table 15 shows what percentage of the total organism/endpoints would be protected assuming a cleanup of sediments occurred to that concentration and all concentration greater than that concentration were removed. The estimated LC10 concentrations are somewhat comparable to the 24 % LC50 concentrations which may not be an incongruous relationship. TPAH concentrations that are causing chronic toxicity may also be causing mortality in a small portions of the population (10%).

Table 16 combines the LC10 and 24% LC50 values with the combined TPAH NOC values ordered in increasing concentrations as they relate to each test result as indicated in the middle column. The column to the right in Table 16 shows what percentage of the total organism/endpoints (n = 16) would be protected assuming a cleanup of sediments occurred to that concentration and all concentration greater than that concentration were removed.

Table 17 shows the ordered TPAH NOC concentrations from Table 16 converted into dry weight TPAH concentrations based on a range of total organic carbon concentrations in the sediments.

<u>Consideration of Elutriate Testing Results</u>

The 1998 toxicity testing included exposing fathead minnow fry to serial dilutions of elutriate from the contaminated wood and sand sites. The serial dilutions created increasing TPAH NOC concentrations when ordered from the highest sand dilution to the undiluted wood elutriate as shown in Table 18. The sediment elutriate was created by mixing the contaminated substrate with water in a 1:4 ratio. The water and sediment were mixed and left to settle for a period of time. After settling, the supernatant was decanted off and the supernatant centrifuged for 45 minutes to further settle out any suspended particles. The supernatant from the centrifuge tube was the elutriate used in the exposure study. To get an idea of the approximate strength of the original pore water in the undiluted sediment and the strength in each of the dilutions, the pore water content of the sampled sediments was assumed to be 50%. As shown in Table 18, in the series of dilutions at 100 %, 50 %, 25 %, 12.5 %, and 6.25 % elutriates , the strength of the original pore water in the sediment is reduced to10.7%, 5.4%, 2.7%, 1.3%, and 0.69%, respectively. If the pore water content of the sample was 30 %, the strength of the original pore water would be reduced to 7 %, 3.5 %, 1.75 %, 0.88 %, and 0.43 %, respectively.

The relevancy of the elutriate testing results to the bottom area of the contaminated sediments off the Lakefront property has been discussed in the ERA and in response comments to commentators on the document. The embayment off the Lakefront property has a relatively shallow water depth (average depth approximately 6 ft.) and is subject to high wind and wave actions. These energies will be translated to disturbances of contaminated bottom sediments and releases of PAHs and other coal tar contaminants in various forms from the pore waters, contaminated wood and sand substrates, and from the NAPL coal tar oils at or near the bottom surface. Oils slicks and sheens are commonly seen following high wind and wave events especially those coming from the direction of the Lake. The elutriate testing is an attempt to simulate the contaminants released from the bottom substrates under such conditions. In the dilution series, it is assumed the TPAH concentrations are reduced proportionate to the amount of the dilution based on the concentration in the 100% elutriate. While the TPAHs are not measured in the elutriate water, the assumption is that they are related to the amount of dilution of the 100% sediment elutriate.

Table 18 shows the increasing effects on the survival and growth endpoints for fathead minnow as the estimated TPAH NOC concentrations increase across the sand and wood substrates. Based on combining the results from the sand and wood substrate serial dilutions, the NOAEC for TPAH NOC is 37 ug/g and the LOAEC is 73 ug/g. These values were not included in the above or below considerations to develop protective levels using various metrics. If they were, they would have supported lower effects threshold concentrations to ensure protection of all organism/endpoints. The elutriate results are important as they relate to the potential impact to early life stages of fish from bottom disturbances of the coal tar-contaminated bottom areas and release of the associated contaminants. The releases can take place from high energy wind and wave events and from diffusion and advection from the substrates and pore water to the overlying surface waters during calmer conditions. At the low end contaminant concentrations, it appears the release of as little as 1% strength pore water can cause significant reductions in growth and biomass and the release as little as 5% strength pore water can cause a significant reduction in survival of fish fry.

This elutriate as used may not represent a worst case situation when suspension of the contaminated sands and silts, wood substrates and coal tar oils and residuals occurs at the site due to wind and surface and internal wave action and seiche effects. The strength, duration, and direction of the wind controls the amount of suspension and the height of the benthic nepheloid layer off the bottom areas of the lake. The latter has been identified as a persistent and particle rich zone of variable suspended solids concentrations that extends up from the bottom areas of Lake Superior. A worst case scenario would be a high energy event that disturbs, releases, and suspends coal tar contaminants from the bottom materials to the overlying water column making them available to aquatic organisms. The organisms would accumulate the PAHs within their tissues. Under calmer conditions which would allow clearing of the water and penetration of UV light deeper into the water column, the UV light would impinge on the organisms, photoactivating certain accumulated PAHs and enhancing their toxicity to the organisms.

The bioassay testing using the elutriate as prepared above is a tool that represents our best estimate to stimulate the dynamic conditions that occur at the site where bottom disturbances, mixing, suspension, release of coal tar contaminants from free oils and residuals, solubilization, redeposition, diffusion, advection etc. occur.

As to how long receptors such as fish eggs, embryos, larvae and young fish are exposed to the coal tar contaminants that are dissolved, associated with suspended particulates, or in released forms of the oils would depend on the disturbance event (e.g. wind direction and speed and wave energy generated), the area of bottom contaminated sediments disturbed, and the life habits of the receptors (e.g. how mobile they are, area use factor, and what level of the water column do they occupy. Larval fish will be confined to relatively small near bottom areas). The sediment dispersion/settling tests reported on in the 1998 ERA (Table 9) indicate that after 11 days, the TPAH concentration in the water column was 4,237 ug/L after an initial concentration of 5,536 ug/L at test initiation. Additionally, there are continual low level releases of coal tar contaminants from the substrates to the overlying water column from advection and diffusion. Using a partitioning model that predicts the pore water concentration of individual PAH compounds based on the organic carbon content and the partitioning coefficient of each compound, the total PAH concentrations were estimated based on past sampling at the site. The estimated TPAH concentrations in the pore water over the site averaged 4,000 ug/L and ranged up to 27,000 ug/L. The estimated concentrations did not consider the solubility limits of the individual PAHs making these estimated pore water concentrations somewhat high. If overestimated by an order of magnitude, the concentrations would still be high and significant.

The following discussion from Petersen et. al (1998) shows the importance of the exposure of early life stages of fish to PAHs: Early life stages of fish are considered to be the most sensitive life stages. The lipid content in the early life stages is higher than in the juvenile adult stages. PAHs will primarily accumulate in the yolk sack lipids which may act as a toxicant sink during the early the embryonic and larval stages. During development of the larvae, PAHs sequestered in the yolk are transported to sensitive organs in which toxic action or metabolism to toxic reactive intermediates may occur. Due to the higher gill surface (and surface in general) to weight ratio in larvae, time to steady state and thus to a “toxic dose equilibrium” is reached sooner in early life stages than in juvenile/adult stages

Recent studies in Alaska (Heintz et al. 1999; Carls et al. 1999) where the fish eggs, embryos, and larvae were exposed to PAHs released from deposited oils in stream bottoms showed that the lowest observed effect concentrations (LOECs) ranged from 0.4 to 1.0 ug/l depending on the species. The LOEC values were based on sublethal responses which included malformations, genetic damage, decreased size, and inhibited swimming that lead to mortality. Increased mortality to salmon embryos occurred when they were exposed to initial aqueousTPAH concentrations of 1.0 ug/L. By inference, the immature life stages of other fish species may be similarly sensitive to low levels PAH exposures from dissolved PAHs in the water column. Enhanced toxicity of PAHs as a result of photoactivation by UV light has been well documented but the experimental setup did not allow for the activation of a significant portion of the PAH molecules. However, photoactivation at the site after the oil spill was likely. Thus the lowest observed effects concentrations in the study that were measured (0.4 ug/L) may actually be conservative compared to the actual on-site conditions due to the spill.

Bottom areas that have relatively unweathered oil associated with them may act as toxic reservoirs that may persist for years until dispersed by a disturbance event. Thus, long-term effects resulting directly from oil exposure are long term in the sense that the PAHs leach over time scales measured in generations.

Heintz et al. (1999) made the following observation:

> “The adverse effect found for embryos exposed to low part-per-million TPAH concentrations reported here by Carls et al. (1999) suggests that restoration of habitats chronically polluted with PAHs may be even more difficult than previously appreciated. The larger more toxic PAHs will most likely persist longest at locations where PAHs are continually leached into receiving aquatic habitats.

The effects of these PAHs on organisms in these habitats may be sublethal at early life stages but may lead to mortality later in life by increasing the vulnerability of these organisms to disease, parasitism, or predation. In our experiments, embryos exposed to PAHs exhibited a variety of adverse effects, and although the frequencies of these effects were often low, the cumulative impact on the exposed populations may be substantial. ”

Table 6 of the 2001 Supplemental ERA shows that 20 species of fish spawn in the Ashland Harbor and 29 species use the harbor for rearing purposes. The habitat of the bay off the Lakefront property is likely used by a number of these fish species during various stages of reproduction and development. The area is important for contributing to the fish populations of Lake Superior.

Summary of TPAH Concentrations Associated With Methods To Derive Protectiveness Levels For Organisms/Endpoints

The five different methods to derive organism/endpoint protective levels discussed above (ANPV, all NOAECs, lowest NOAECs, LC10 + 24%LC50, and MATC) are summarized below. The summary values are based on a desired protectiveness value of 80% or greater. Protectiveness values that bracket the 80% value from the various methods are shown. Based on a combining of all the values from the table below, an 80 % protectiveness level for all the organism/endpoints would be achieved at TPAH NOC concentrations in the 100 – 200 ug/g range and an associated dry weight concentration range of 2 to 28 mg/kg depending on the TOC content of the substrate involved. Given the uncertainties of the presence of coal tars acting as a possible TOC source at higher TOC concentrations and the MATC value not defining the lower threshold effect bounds, it is recommended that the dry weight TPAH protective range be from 2 to 15 or 20 mg/kg. Organism/ endpoints that would not fall within this protective range means there is the possibility that the endpoints of growth or survival would be adversely effected.

Method Used to Derive Protectiveness Levels For Organism/Endpoints	TPAH NOC Concentration (ug/g OC) Associated With Protectiveness Levels Bracketing 80%		TPAH NOC Concentrations Converted to a Range of Dry weight Concentrations (mg/kg) Based on 2% - 12 % TOC Content in Sediments	
	ug/g OC	Protectiveness Level	mg/kg TPAH at 2 % TOC	mg/kg TPAH at 12 % TOC
Adjusted NOAEC Protective Value (ANPV)	**92**	**100**	**1.84**	**11.04**
	104	**60**	**2.04**	**12.84**
Lowest NOAECs	**92**	**100**	**1.84**	**11.04**
	115	**60**	**2.30**	**13.80**
LC 10 + 24% LC50	**74**	**100**	**1.48**	**8.88**
	120	**75**	**2.40**	**14.40**
All NOAECs	**92**	**100**	**1.84**	**11.04**
	115	**74**	**2.30**	**13.80**
MATC	**232**	**100**	**4.64**	**27.84**
	236	**80**	**4.72**	**28.32**

Role of LOAEC Values In the Risk Description Process

The LOAEC values derived from each of the toxicity tests conducted in 1998 and 2001 are shown in Table 19. The survival and growth values are expressed as a percent reduction in survival and/or growth at the sudy sites compared to the reference site. As discussed above, the *lowest observed adverse effect concentration* (LOAEC) is the lowest test concentration evaluated in the site media that has a statistically significant adverse effect on the exposed organisms compared to the mean response at the control site test media or the reference site media in a specific test. In the Ashland Lakefront ERA, a statistically significant result that represented a 20% or more reduction in the growth, survival, or reproductive endpoints in the study site sediments compared to the reference site sediment results were interpolated to have ecological significance to populations in the field. As can be seen in Table 19 below, the reductions in survival and growth associated with the LOAEC values were much greater than the 20% value.

U.S. EPA (1997) indicates that a key output of the risk characterization step is the establishment of contaminant concentrations in the media that bound the thresholds for estimated adverse ecological effects which is reflected in contamination levels posing no or minimal ecological risks to concentrations associated with observations or predictions that adverse ecological impacts could occur. The latter is further defined as the lowest contamination levels identified as likely to produce adverse ecological effects. U.S. EPA goes on to say that one of the ways the upper bound effects threshold can be developed is by using the LOAEC toxicity values. As can be seen from Table 19,
the derived LOAECs from all of the 1998 and 2001 bioassays represent large reductions in survival and growth well beyond the 20% value. The TPAH contamination levels associated with the 20% reduction where adverse ecological effects are predicted to start taking place is at some intermediate concentration between the NOAEC and the LOAEC. To find what the exact concentration is, toxicity testing would have to be performed at concentrations in this intermediate range.

The point is that the LOAEC values derived from the 1998 and 2001 toxicity tests should not be used to define the upper effect threshold bounds if it is implied that the associated TPAH concentrations would be considered acceptable as an upper risk level for use in deriving remediation goals for the site. The LOAEC values go beyond the acceptable risk levels. Ecological impacts at the LOAECs would be predicted to be severe and are unacceptable. Another approach needs to be used to define the upper bound effects threshold concentrations that are associated with the lowest concentration at which statistically significant reductions in the endpoints of 20% start to occur An alternative may be the MATC values that lie between the NOAEC and LOAEC values.

Comparisons Of the Bioassay Results and the Preliminary Remediation Goals Derived From the Results With Sediment Quality Guideline Benchmarks From Other Sources

A number of sediment quality guidelines (SQGs) have been developed for PAHs using empirical and mechanistic approaches (equilibrium partitioning of PAHs from organic carbon to pore water) . The ERA and supplemental ERA compare some of the more currently developed guideline values (e.g., U.S. EPA, 2001 Draft, MacDonald et. al. 2000) with the Ashland bioassay results. While we don't advocate using the sediment guidelines on a stand-alone basis for making remediation decisions for a site, we believe it is appropriate to use the guideline values in the baseline ERA on an integrated and comparative basis with the effect-related concentrations found in the site-specific studies. Comparing the values from the site-specific studies and the SQGs helps to determine if the site-specific effect-concentration relationships are validated by the effect values in the guidelines. There is nothing that precludes this process from being done and it lends a reality check to any results by comparing site results to a much larger toxicity data base for which effect concentrations have been established. If the effect concentrations from any field data set are compatible with the effect concentrations of a much larger data set, it would strengthen the case for the site-specific effect concentrations and any sediment management decisions based on the site data.

The particular set of SQGs used for comparison purposes are discussed in the ERA and compared with the chemical and bioassay result data principally in Tables 17,18, and 19 of the supplemental ERA. The data is summarized in the appendix table below. Some basic information from the two of the guidelines used for comparative purposes are as follows:

1) U.S. EPA.s (2000 Draft) Equilibrium Partitioning Sediment Guidelines for PAH Mixtures For the Protection of Benthic organisms. See Sections 6.1.3 and 6.2.2 of the ERA Supplement (SEH, 2002) for basic information on this approach for developing SGQs for PAHs. The approach is based on a partitioning model that predicts the concentration of each PAH in the sediment pore water based on the organic carbon content of the sediments and the partitioning coefficient of the PAH. The concentration is divided by the chronic toxicity value for that PAH to derive a toxicity unit (TU) value. The toxicity unit values for the individual PAH are summed to yield a Σ PAH ESG TU value. If the summed TU value exceeds 1, sensitive benthic organisms may be affected by chronic toxicity. Based on an acute to chronic ratio of 4.16, if the summed TU value exceeds 4.16, one would expect lethal effects for sensitive species. Between Σ ESH TU of 1 and 4, only chronic effects are expected, unless the species are unusually sensitive. The summed TU values and predicted effects relationships are summarized below.

U.S. EPA Equilibrium Partitioning Sediment Guidelines Summed Toxic Units For Total PAHs		
ESG $\sum TU_{TOTPAHS}$		
> 1.0	1.0 to < 4.16	> 4.16
No chronic toxicity effects	Sensitive benthic organisms may be affected by chronic toxicity. Lethal effects to unusually sensitive species	Acute toxicity. Lethal effects for sensitive species

2) MacDonald et al. (2000) Consensus-Based Sediment Quality Guidelines. MacDonald et. al combined several sets of SQGs for the protection of benthic organisms that used a variety of approaches and derived two effect-related concentrations – a threshold effect concentration (TEC) where the probability of toxicity is low or absent and a probable effect concentration (PEC) where the probability of toxicity is high. For total PAHs, A PEC quotient (PEC-Q) can be derived by dividing the total PAH concentration in the sediments at a site by the PEC value for total PAHs. As the PEC-Q increases, the incidences of toxicity as reflected in reductions in the survival, growth, and reproduction endpoints is predicted to increase. In assessing the degree of concordance that exists between the effect-related chemical concentrations and the incidence of toxicity, it has been demonstrated that for the most reliable consensus-based SQG contaminants, there is a consistent and incremental increase in the incidence of toxicity to sediment dwelling organisms with increasing chemical concentrations.

A relative idea of the of the relationship between the range of PEC quotients for total PAHs and the percent incidences of toxicity found in bioassays at the PEC-Q values can be seen in the following two sources.

MacDonald et al. (2000) also looked at the predictive ability of the CBSQGs. To examine the relationships between the degree of chemical contamination and probability of observing toxicity in freshwater sediments, the incidence of toxicity within various ranges of mean PEC quotients was calculated from an existing database. The data were plotted in a graph (Table 1, MacDonald et al. 2000). The interpolated data from this graph is in the table below. MacDonald et al. found that subsequent curve-fitting indicated that the mean PEC-quotient is highly correlated with incidence of toxicity ($r^2 = 0.98$), with the relationship being an exponential function. The resulting equation ($Y = 101.48 (1-0.36^X)$ can be used to estimate the probability of observing sediment toxicity at any mean PEC quotient.

Relationship between Mean PEC Quotient and Incidence of Toxicity in Freshwater Sediments (Derived from MacDonald et al. 2000a)	
Mean PEC Quotient	**Average Incidence of Toxicity**
0	0
0.25	20
0.50	40
0.75	54
1.00	64
1.25	70
1.50	77
1.75	84
2.00	87
2.25	90
2.50	92
2.75	95
3.00	96
3.25	98
3.50	99
3.75	99.5
≥ 4.00	100

The database used by Ingersoll et al. (2001) to determine the ability of the PEC-Qs to predict toxicity is based on testing freshwater sediments from a number of sites using 10- to 42-day toxicity tests with the amphipod *Hyalella azteca* or the 10- to 14-day toxicity tests with the midges *Chironomus tentans* or *C. riparius.* Toxicity of samples was determined as a significant reduction in survival or growth of the test organisms relative to a control or reference sediment. A relative idea of the predictive ability of the overall mean PEC-Qs and individual PEC-Qs for each group of chemicals is shown in the table below from Ingersoll et al. (2001). Mean PEC quotients were calculated to provide an overall measure of chemical contamination and to support an evaluation of the combined effects of multiple contaminants in sediments.

Test Species and Test Duration	Incidence of Toxicity (% of samples where toxicity observed versus no toxicity).					Total No. of Samples
	Range of PEC-Q $_{\text{total PAHs}}$					
	< 0.1	0.1 to < 0.5	0.5 to < 1.0	1.0 to < 5.0	> 5.0	
Hyalella azteca 10 to 14 day tests	25 (123)	33 (76)	35 (20)	49 (33)	100 (14)	266
Hyalella azteca 28 to 42 day tests	8 (57)	64 (37)	55 (9)	NC	100 (6)	109
Chironomus spp. 10 to 14 day tests	26 (64)	33 (73)	77 (13)	85 (20)	71 (7)	177

The relative relationship between the ESG TU and PEC-Q benchmarks along with the results of the toxicity testing and the metrics (e.g., LC 10, LC 50, and 24% LC 50) from these results can be seen in the Appendix Table at the end of this report. Table 19 below relates the protectiveness levels derived from combining the LC 10 and chronic toxicity values (24% LC50) and the toxicity benchmarks derived from the gradient of NOC TPAH site concentrations. The protectiveness level of approximately 80% related to chronic toxicity and high survivability is related to a ESG $TU_{TOTPAHs}$ of 2 which is associated with chronic toxicity to some sensitive benthic organisms. The gradient of PEC-Q values does not increase in as orderly a manner as site NOC TPAH concentrations increase compared to the $ESGTU_{TOTPAH}$ values. Based on a reasonable conservative interpolation of the PEC-Q values, PEC-Q values of 0.1 to 0.2 are associated with an 80% level of effectiveness.

Tables Referenced In the Above Text

Table 1. Derivation if Maximum Acceptable Toxicant Concentration (MATC)[1.] From the Combined 1998 and 2001 Bioassays.		
TPAH NOC ug/g OC	**Organism/Endpoint (n = 15)**	**% of Organism/Endpoints Protected**
232	DM-48 UV; CT-10; HA-10	100 %
236	HA-28 UV; PP-7	80 %
713	CT-10	67 %
1,582	LV-10 UV; LV-10; HA-10; CT-10	60 %
1,714	HA-28 UV; CT-10	33 %
2,102	HA-28	20 %
4,399	HA-28	13 %
6,249	PP-7	7 %
1. MATC derived by calculating the geometric mean (U.S. EPA, 1998) of the NOAEC and LOAEC values based on the bioassay results for each organism tested.		

Table 2. Conversion of TPAH Concentrations Related to MATCs To a Range of Dry Weight Bulk sediment Concentrations									
MATC TPAH ug/g NOC	TPAH Dry Weight Bulk Sediment Concentrations At Increasing Total Organic Carbon Percentages in the Sediment **mg/kg**								**Est. % Organism/ Endpoint Protected At Given TPAH Concentration**
	1%	2 %	4%	6 %	8 %	10 %	12%	14 %	
232	2.32	4.64	9.28	13.92	18.56	23.20	27.84	32.48	100 %
236	2.36	4.72	9.44	14.16	18.88	23.60	28.32	33.04	80 %
713	7.13	14.26	28.52	42.78	57.04	71.30	85.56	99.82	67 %
1,582	15.82	31.64	63.28	94.92	126.6	158.20	189.8	221.5	60 %
1,714	17.14	34.28	68.56	102.8	137.1	171.4	205.7	239.9	33 %
2,102	21.02	42.04	84.08	126.1	168.2	210.2	252.2	294.3	20 %
4,399	43.99	87.98	175.9	263.9	351.9	439.9	527.9	615.9	13 %
6,249	62.49	124.9	249.9	374.9	499.9	624.9	749.9	874.9	7 %

Table 3. Derived NOAEC and LOAEC Values Based on Organic Carbon TPAH Concentrations In the 2001 Bioassays					
Wood Substrate					
TPAH Concentrations In Test Substrate		Bioassay Tests			
NOC ug/g OC	**Dry Weight mg/kg**	**HA-28 UV**	**HA-28**	**CT-10**	**PP-7**
66 Ref.	12.6 Ref.				
105		**ANPV**			**ANPV**
143	20.1	NOAEC			NOAEC
271				**ANPV**	
399	103.8	LOAEC		NOAEC	LOAEC
855			**ANPV**		
1,310	301		NOAEC	LOAEC	
1,582	759		LOAEC		
1,963	104				
2,873	661				
Sand Substrate					
148 Ref.	3.1 Ref.				
442		**ANPV**	**ANPV**	**ANPV**	
735	16.2	NOAEC	NOAEC	NOAEC	
2,366					**ANPV**
3,996	80	LOAEC		LOAEC	NOAEC
4,842	823		LOAEC		LOAEC
9,978	249				
23,231	836				
123,568	235				

Table 4. Lowest NOAEC and LOAEC Values In the Data Set From the Combined 2001 Bioassays Ran On the Sand and Wood Substrates					
TPAH Concentrations In Test Substrate		Bioassay Tests			
NOC ug/g OC	**Dry Weight mg/kg**	**HA-28 UV**	**HA-28**	**CT-10**	**PP-7**
66 W Ref	12.6				
105		**ANPV**			**ANPV**
143 W	20.1	NOAEC			NOAEC
148 S Ref	3.1	LOAEC			
274				**ANPV**	
399 W	103.8			NOAEC	LOAEC
567			**ANPV**		
735 S	16.2		NOAEC		
1,310 W	301		LOAEC	LOAEC	
1,582 W	759				
1,963 W	104				
2,873 W	661				
3,996 S	89				
4,842 S	823				
9,978 S	249				
23,231 S	836				
123,568 S	235				

Table 5. Lowest NOAEC and LOAEC Values In the Data Set From the Combined 1998 Bioassays Ran On the Sand and Wood Substrate							
TPAH Concentrations In Test Substrate		**Bioassay Tests**					
NOC ug/g OC	**Dry Weight mg/kg**	**HA-10**	**CT-10**	**LV-10**	**LV-10 UV**	**DM-48**	**DM-48 UV**
92		**ANPV**	**ANPV**	**ANPV**	**ANPV**		**ANPV**
92 SR	0.42	NOAEC	NOAEC	NOAEC	LOAEC		NOAEC
104							
115WR	6.5						
350					Adverse Effects For All	**ANPV**	
584 S	1.5	LOAEC	LOAEC	LOAEC		NOAEC	LOAEC
21,776 W	370					No Adverse Effects	

Table 6. Lowest NOAEC Values From the Combined 1998 and 2001 Bioassays Based On NOC TPAH Concentrations Compared to Percentages of Organism/Endpoints protected			
TPAH ug/g OC	**NOAEC Concentration**	**Test Organism and Endpoint**	**Estimated % of Organism/Endpoints Protected At the Given TPAH NOC Concentration**
66 W Ref			
92	NOAEC	CT-10; LV-10 UV; DM-48 UV; HA-10; LV-10	100 %
92 S Ref			
104			
105			
115 W Ref			
143 W	NOAEC	HA-28 UV; PP-7	50%
148 S Ref			
274			
350			
399 W	NOAEC	CT-10	30 %
567			
584 S	NOAEC	DM-48	20 %
735 S	NOAEC	HA-28	10 %
1,310 W			0%
1,582 W			0%
1,963 W			0%
2,873 W			0%
3,996 S			0%
4,842 S			0%
9,978 S			0%
21,776 W			0%
23,231 S			0%
123,568 S			0%

Table 7. Adjusted Lowest NOAEC Values From the Combined 1998 and 2001 Bioassays Based On NOC TPAH Concentrations Compared to Percentages of Organism/Endpoints protected			
TPAH ug/g OC	**Adjusted NOAEC Concentration**	**Test Organism and Endpoint**	**Estimated % of Organism/Endpoints Protected At the Given TPAH NOC Concentration**
66 W Ref			
92	ANPV	CT-10; LV-10UV; DM-48 UV; HA-10; LV-10	100 %
92 S Ref			
105	ANPV	HA-28 UV; PP-7	50%
115 W Ref			
143 W			
148 S Ref			
274	ANPV	CT-10	30 %
350	ANPV	DM-48	20 %
399 W			
567	ANPV	HA-28	10 %
584 S			0%
735 S			0%
1,310 W			0%
1,582 W			0%
1,963 W			0%
2,873 W			0%
3,996 S			0%
4,842 S			0%
9,978 S			0%
21,776 W			0%
23,231 S			0%
123,568 S			0%

Table 8. Conversion of All 1998 and 2001 TPAH NOC NOAEC Concentrations To A Range of Dry Weight Bulk Sediment Concentrations Related to Percentages of Organism/Endpoints Protected													
NOAEC Concentration TPAH ug/g NOC	**TPAH Dry Weight Bulk Sediment Concentration At Increasing Total Organic Carbon Percentages In the Sediment mg/kg**												**Estimated % of Organisms/ Endpoints Protected At the Given TPAH Concentration**
	1 %	2 %	3 %	4 %	5 %	6 %	7 %	8 %	9 %	10 %	12 %	15 %	
92	0.92	1.84	2.76	3.68	4.60	5.52	6.44	7.36	8.28	8.28	11.04	13.80	100 %
115	1.15	2.30	3.45	4.60	5.75	6.90	8.05	9.20	10.35	10.35	13.80	17.25	74 %
143	1.43	2.86	4.29	5.72	7.15	8.58	10.01	11.44	12.87	12.87	17.16	21.45	58 %
399	3.99	7.98	11.97	15.96	19.95	23.94	27.93	31.92	35.91	35.91	47.88	59.85	47 %
584	5.84	11.68	17.52	23.36	29.20	35.04	40.88	46.72	52.56	52.56	70.08	87.60	42 %
735	7.35	14.70	22.05	29.40	36.75	44.10	51.45	58.80	66.15	66.15	88.20	110.3	32 %
1,310	13.10	26.20	39.30	52.40	65.50	78.60	91.70	104.8	117.9	117.9	157.2	196.5	16 %
3,996	40.0	80.00	120.0	160.0	200.0	240.0	280.0	320.0	360.0	400.0	480.0	600.0	11%
21,779	218.0	436	654	1090	1090	1308	1526	1744	1962	2180	2616	3270	5 %

Table 9. Conversion of Lowest 1998 and 2001 TPAH NOC NOAEC Concentrations To A Range of Dry Weight Bulk Sediment Concentrations Related to Percentages of Organism/ Endpoints Protected													
NOAEC Concentration TPAH ug/g NOC	TPAH Dry Weight Bulk Sediment Concentration At Increasing Total Organic Carbon Percentages In the Sediment mg/kg												Estimated % of Organisms/ Endpoints Protected At the Given TPAH Concentration
	1 %	2 %	3 %	4 %	5 %	6 %	7 %	8 %	9 %	10 %	12 %	15 %	
92	0.92	1.84	2.76	3.68	4.60	5.52	6.44	7.37	8.28	9.20	11.04	13.80	100 %
115	1.15	2.30	3.45	4.60	5.75	6.90	8.05	9.20	10.35	11.5	13.80	17.25	60 %
143	1.43	2.86	8.58	5.72	7.15	8.58	10.01	11.44	12.87	14.30	17.16	21.45	50 %
399	3.99	7.98	11.97	15.96	19.95	23.94	27.93	31.92	35.91	39.90	47.88	59.85	30 %
584	5.84	11.68	17.52	23.36	29.20	35.04	40.88	46.72	52.56	58.40	70.08	87.60	20 %
735	7.35	14.7	22.05	29.40	36.75	44.10	51.45	58.80	66.15	73.50	88.20	110.3	10 %

Table 10. Conversion of Lowest TPAH NOC Adjusted NOAEC Concentrations To A Range of Dry Weight Bulk Sediment Concentrations													
Adjusted NOAEC Concentration TPAH ug/g NOC	TPAH Dry Weight Bulk Sediment Concentration At Increasing Total Organic Carbon Percentages In the Sediment mg/kg												Estimated % of Organisms/ Endpoints Protected At the Given TPAH Concentration
	1 %	2 %	3 %	4 %	5 %	6 %	7 %	8 %	9 %	10 %	12 %	15 %	
92	0.92	1.84	2.76	3.68	4.60	5.52	6.44	7.37	8.28	9.20	11.04	13.80	100 %
104	1.04	2.04	3.12	4.16	5.20	6.24	7.28	8.32	9.36	10.40	12.48	15.60	60 %
105	1.05	2.10	3.15	4.20	5.25	6.30	7.35	8.40	9.45	10.50	12.60	15.75	50 %
274	2.74	5.48	8.22	10.96	13.7	16.44	19.18	21.92	24.66	27.40	32.88	41.10	30 %
350	3.50	7.00	10.5	14.00	17.5	21.00	24.50	28.00	31.5	35.00	42.00	52.5	20 %
587	5.87	11.74	17.61	23.48	29.35	35.22	41.09	46.96	52.83	58.70	70.44	88.05	10 %

Table 11 . Adjustment of LC50 TPAH Concentrations (From 2001 Bioassays to Estimated Chronic Toxicity Levels For TPAH Concentrations

	Calculated TPAH 10-d LC50 Values				24% of TPAH LC50 Values to Estimate Chronic Toxicity Value			
TPAH	Sand Substrate							
Sand	HA-28 UV	HA-28	CT-10	PP-7	HA-28 UV	HA-28	CT-10	PP-7
NOC ug/g OC	2,671	9,861	3,144	8,264	642	2,370	756	1,987
Dry Wt. mg/kg	94	354	111	296	23	85	27	71
Wood Substrate								
TPAH	HA-28 UV	HA-28	CT-10	PP-7	HA-28 UV	HA-28	CT-10	PP-7
NOC ug/g OC	307	947	777	1,165	74	228	187	280
Dry Wt. mg/kg	135	460	373	570	33	111	90	14

Table 12. 24% LC50 (Chronic Toxicity Value) Concentrations Related to Percentages of Organisms/Endpoints Protected.

TPAH NOC ug/g OC	Organism/Endpoint	Estimated % of Organism/Endpoints Protected at the Given TPAH NOC Concentration
74	HA-28 UV	100 %
187	CT-10	88 %
228	HA-28	75 %
280	PP-7	63 %
642	HA-28 UV	50 %
756	CT-10	38 %
1987	PP-7	25 %
2370	HA-28	13 %

Table 13. Estimated LC_{10} Concentrations

Sand	HA-28 UV	HA-28	CT-10	PP-7
NOC ug/g OC	700	1,960	700	1,400
Dry Wt. mg/kg	25	70	25	50
Wood				
NOC ug/g OC	80	120	100	100
Dry Wt. mg/kg	40	60	50	50

Table 14. Estimated LC10 Concentrations Related to Percentages of Organisms/Endpoints Protected.		
TPAH NOC ug/g OC	**Organism/Endpoint**	**Estimated % of Organism/Endpoints Protected at the Given TPAH NOC Concentration**
80	HA-28 UV	100 %
100	CT-10; PP-7	88 %
120	HA-28	63 %
700	HA-28 UV; CT-10	50 %
1,400	PP-7	25 %
1,960	HA-28	13 %

Table 15. Combined LC10 and 24% LC50 Values (Tables 13 and 15 above) Related to Percentages of Organism/Endpoints Protected.		
TPAH NOC ug/g OC	**Organism/Endpoint**	**Estimated % of Organism/Endpoints Protected at the Given TPAH NOC Concentration**
74	HA-28 UV 24% LC50	100 %
80	HA-28 UV LC 10	94 %
100	CT-10 LC10; PP-7 LC10	88 %
120	HA-28 LC10	75 %
187	CT-10 24% LC50	69 %
228	HA-28 24% LC50	63 %
280	PP-7 24% LC50	56 %
642	HA-28 UV 24% LC50	50 %
700	HA-28 UV LC10; CT-10 LC10	44 %
756	CT-10 24% LC50	31 %
1,400	PP-7 LC10	25 %
1,960	HA-28 LC10	19 %
1,987	PP-7 24% LC50	13 %
2,370	HA-28 24% LC50	6 %

Table 16. Conversion of TPAH NOC Concentrations Related to Combined 24% LC50 and LC10 Values To a Range of Dry Weight Bulk Sediment Concentrations									
LC10 and 24% LC50 TPAH ug/g NOC	TPAH Dry Weight Bulk Sediment Concentrations At Increasing Total Organic Carbon Percentages in the Sediment mg/kg								Est. % Organism/ Endpoint Protected At Given TPAH Concentration
	1%	2 %	4%	6 %	8 %	10 %	12%	14 %	
74	0.74	1.48	2.96	4.44	5.92	7.40	8.88	10.36	100 %
80	0.80	1.60	3.20	4.80	6.40	8.00	9.60	11.20	94 %
100	1.00	2.00	4.00	6.00	8.00	10.00	12.00	14.00	88 %
120	1.20	2.40	4.80	7.20	9.60	12.00	14.40	16.80	75 %
187	1.87	3.74	7.48	11.22	14.96	18.70	22.44	26.18	69 %
228	2.28	4.56	9.12	13.68	18.24	22.80	27.36	31.92	63 %
280	2.80	5.60	11.20	16.80	22.40	28.00	33.60	39.20	56 %
642	6.42	12.84	25.68	38.52	51.36	64.20	77.04	89.88	50 %
700	7.00	14.00	28.00	42.00	56.00	70.00	84.00	98.00	44 %
756	7.56	15.12	30.24	45.36	60.48	75.60	90.72	105.8	31 %
1,400	14.00	28.00	56.00	84.00	112.0	140.0	168.0	196.0	25 %
1,960	19.60	39.20	78.40	117.6	156.8	196.0	235.2	274.0	19 %
1,987	19.87	39.74	79.48	119.2	158.9	198.7	238.4	278.0	13 %
2,370	23.70	47.40	94.8	14.22	189.6	237.0	284.4	331.8	6 %

Table 17 . Results of *Pimephales promelas* (Fathead minnow) Exposure to Serial Dilutions of Elutriate From Contaminated Sand and Sediment Sites in 1998 (Taken from the above Table ordering of organic carbon normalized site concentrations NOC)					
Contaminated Sand Substrate				**Test Endpoints**	
TPAH NOC ug/g OC	**Dry Wt. mg/kg**	**% Sediment Elutriate**	**Pore Water Strength Est. Based on 50% Water by Volume in Sediments**	**% Survival**	**% Reduction in Weight Compared to Reference Site**
37 SE	0.091	6.25	0.69	95	-19
73 SE	0.18	12.5	1.33	85	-42*
146 SE	0.37	25	2.67	78	-67*
292 SE	0.73	50	5.35	38*	-79*
584 SE	1.5	100	10.7	5*	-68*
Contaminated Wood Substrate				**Test Endpoints**	
1,361 WE	23.1	6.25	0.69	18*	-66*
2,722 WE	46.3	12.5	1.33	0*	-100*
5,444 WE	92.6	25	2.67	0*	-100*
10,888 WE	185	50	5.35	0*	-100*
21,776 WE	370	100	10.7	0*	-100*

Table 18. Degree of Effects Associated With the LOAEC Value Derived From the 1998 and 2001 Bioassay Results.						
Test Organism	**Sand Substrate**			**Wood Substrate**		
	LOAEC TPAH ug/g OC	**Test Endpoint: % Reduction Compared to Reference Site**		**LOAEC TPAH ug/g OC**	**Test Endpoint: % Reduction Compared to Reference Site**	
		Survival	**Growth**		**Survival**	**Growth**
1998 Bioassays						
DM-48 UV	**584**	**No Survival**	**- 100 %**	**115**	**- 92 %**	**NA**
DM-48	**No AE**	**---**	**---**	**No AE**	**---**	**---**
LV-10 UV	**92**	**- 60%**	**NA**	**115**	**No Survival**	**- 100 %**
LV-10	**584**	**- 27 % N.S.[1.]**	**- 61 %**	**21,776**	**- 85 %**	**- 93 %**
HA-10	**584**	**- 7 % N.S.**	**- 42 %**	**21,776**	**- 73 %**	**- 86 %**
CT-10	**584**	**- 13 % N.S.**	**- 23 %**	**21,776**	**- 28 %**	**- 47 %**
2001 Bioassays						
HA-28 UV	**3,996**	**- 26 %**	**See Note 2.**	**399**	**- 63 %**	**See Note 2.**
HA-28	**4,842**	**- 46 %**	**See Note 2.**	**1,582**	**- 97 %**	**See Note 2.**
CT-10	**3,996**	**- 76 %**	**See Note 2.**	**1,310**	**No Reduction**	**- 30 %**
PP-7	**4,842**	**- 87 %**	**- 68 %**	**399**	**No Reduction**	**- 25 %**

Notes:

1. N.S. Indicates reduction is not statistically significant.

2. While mean weight differences expressed as mg/ individual surviving organism at the study sites are not significantly different than the mean weights of the individual organisms at the references sites, the overall reduction in survival of the organisms means that total biomass will be significantly reduced. For example, in the 2001 CT-10 bioassay on the sand substrate, there was an average of 2.3 surviving organisms in the 8 replicates from the study site at the end of the test versus an average of 9.5 in the replicates from the reference site. While there was not statistically significant difference in the mean weight of the survivors, there was a significant difference in overall biomass (2.5 survivors x 0.58 mg/organism = 3.63 mg of total biomass at the study site versus 9.5 survivors x 0.83 mg/organism = 7.89 mg.

Table 19. Relationships between Protectiveness Levels Based on LC 10 and 24% LC50 (Chronic Toxicity) and Benchmark Toxicity Values.			
TPAH NOC ug/g OC	**Protectiveness Levels of Organisms and Endpoints (See Table 15 above)**	**ESG $TU_{TOTPAHs}$**	**PEC-Q**
66		0.8	0.4
74	100 %		
80	94 %		
92		1	0.01
100	88 %		
115		2	0.2
120	75 %		
143		2	0.6
148		2	0.09
187	69 %		
228	63 %		
280	56 %		
399		6	3.6
584		9	0.06
642	50%		
700	44 %		
735		11	0.6
756	31 %		
1,310		19	10.9
1,400	25 %		
1,582		23	28
1,960	19 %		
1,963		28	3.7
1,987	13 %		
2,370	6 %		
2,873	0 %	42	24
3,996	0 %	62	3
4,842	0 %	76	30
9,978	0 %	156	9
21,776	0 %	323	13
23,231	0 %	369	32
123,568	0 %	1,908	9

References

Barnthouse, L.W. 1993. Population level effects. In: Ecological Risk Assessment. Edited by G. W. Suter II. Lewis Publishers. Ann Arbor, MI.

Burton, G.A., Jr. 1991. Annual Review. Assessing the toxicity of freshwater sediments. Environ. Toxicol. Chem. 10:1585-1627.

Carls, M.G., S.D. Rice, and J.E. Hose. 1999. Sensitivity of fish embryos to weathered crude oil: Part I. Low-level exposure during incubation causes malformations, genetic damage, and mortality in larval Pacific herring (*Clurea Pallasi*). Environ. Toxicol. Chem. 18:481-493.

Chapman P.F., M. Crane, J. Wiles, F. Noppert, and E. McIndoe. 1996. Improving the quality of statistics in regulatory ecotoxicity tests. Ecotoxicology. 5:169-186.

Chapman, P.M., M. Cano, A.T Fritz, C. Gaudet, C.A. Menzie, M. Sprenger, and W.A. Stubblefield. 1997. Critical issues in methodological uncertainty. Session 4. Contaminated site cleanup decisions. Chapter 7. Workgroup summary report on contaminated site cleanup decisions. Proceedings of the Pellston Workgroup on Sediment Ecological Risk Assessment. April 23-28, 1995. Pacific Grove, CA. Edited by C.G. Ingersoll, T. Dillon, and G.R. Biddinger. Socciety of Environmental Toxicology and Chemistry. SETAC Press.

Crane, M. and M.C. Newman. 2000. What level of effect is a no observed effect? Environ. Toxicol. Chem. 19;516-519.

Di Torro, D.M. et al. 1991. Technical basis for establishing sediment quality criteria for nonionic organic chemicals using equilibrium partitioning. Environ. Toxicol. Chem. 10:1541-1583.

Duke, L.D, M. Taggart. 2000. Uncertainty factors in screening ecological risk assessments. Environ. Toxicol. Chem. 19:1668-1680.

Giesy, J.P., C.J. Rosiu, and R.L. Graney. 1990. Hazard Assessment. Benthic invertebrate bioassays with toxic sediment nd pore water. Environ. Toxicol. Chem. 9:233-248.

Heintz, R.A., J.W. Short, and S.D. Rice. 1999. Sensitivity of fish embryos to weathered crude oil: Part II. Increased mortality of pink salmon (*Oncorhynchus Gorbuscha*) embryos incubating downstream from weathered Exxon Valdez crude oil. Environ. Toxicol. Chem. 18:494-503.

Ingersoll, C.G., G.T. Ankley, R. Baudo, G.A. Burton, W. Lick, S.N. Luoma, D.D. MacDonald, T.B. Reynoldson, K.R. Solomon, R.C. Swartz, and W.J. Warren-Hicks. 1997. Chapter 18. Workgroup summary report on uncertainty evaluation of measurement endpoints used in sediment ecological risk assessment. Proceedings of the Pellston Workshop on Sediment Ecological Risk Assessment. April 23-28, 1995. Edited by C.G. Ingersoll, T. Dillon, G.R. Biddinger. Society of Environmmmmental Toxicology and Chemistry (SETAC) Press.

Ingersoll, C.G., D.D. MacDonald, N. Wang, J.L. Crane, L.J. Field, P.S. Haverland, N.E. Kimble, R.A. Lindskoog, C. Severn, and D.E. Smorong. 2001. Predictions of sediment toxicity using consensus-based freshwater sediment guidelines. Arch. Environ. Contam. Toxicol. 41:8-21.

Lawrence, G. 1999. EC20 determinations for toxicity tests in aquatic risk assessments. SETAC News. July 1998.

LSRI (Lake Superior Research Institute (LSRI). 2001. Evaluation of aquatic sediments from the Ashland Harbor NSP Lakefront Site for toxicity to *Chironomus tentans, Pimephales promelas, and Hyalella azteca.* By M. Balcer, M. TenEyck, T. Markee, and D.L. Brooke. LSRI. University of Wisconsin-Superior. Superior, WI. Final Report.

MacDonald, D.D., C.G. Ingersoll, and T.A. Berger. 2000. Development and evaluation of consensus-based sediment quality guidelines for freshwater ecosystems. Arch. Environ. Contam. Toxicol. 39:20-31.

Michelsen, T. 1999. Contaminated sediments: When is cleanup required? The Washington State approach. Pgss. 74-77. *In* Deciding when to intervene. Data interpretation tools for making sediment management decisions beyond source control. Sediment Priority Action Committee. Great Lakes Water Quality Board. Report to the International Joint Commission. Based on a workshop at the Great Lakes Institute for Environmental Research in Windsor, Ontario. December 1-2, 1998.

Nelson, M.K., P.F. Landrum, G.A. Burton, Jr., S.J. Klaine, E.A. Crecelius, T.D. Byl, D.C. Gossiaux, V.N. Tsymbal, L. Clevland, C.G. Ingersoll, and G. Sasson-Brickson. 1993. Toxicity of contaminanted sediments in dilution series with control sediments. Chemosphere. 9:1789-1812.

Newman, M.C. 1995. Quantitative Methods in Aquatic Ecotoxicology. CRC, Chelsea, MI, USA.

Newman, M.C., D.R. Ownby, L.C.A. Mezin, D.C. Powell, R.L. Tyler R.L. Christensen, S.B. Lerberg, and B-A Anderson. 2000. Applying species-sensitivity distributions in ecological risk assessment: Assumptions of distribution type and sufficient numbers of species. Environ. Toxicol. Chem. 2:508-515.

Odum, E.P. 1966. Fundamentals of Ecology. W.B. Saunders Company. Philadelphia, PA.

Petersen, G.I. and P. Kristensen. 1998. Bioaccumulation of lipophilic substances in fish early life stages. Environ. Toxicol. Chem. 17:2831-2843.

Rand, G.M. and S.R. Petrocelli. 1985. Fundamentals of Aquatic Toxicology. Methods and Applications. Edited by B.L. Donley and M. Dorfman. Hemisphere Publishing Corp.

Sprague, . . 1971.

Suter, G.W. II. 1996. Risk characterization for ecological risk assessment of contaminated sites. Oak Ridge National Laboratoty. Oak Ridge, TN. ES/ER/TM-200.

Suter, G.W. II and C.L. Tsao. 1996. Toxicological benvhmarks for screening potential contamination for effects on aquatic biota: 1996 revision. Oak Ridge National Laboratory. Oak Ridge, TN. ES/ER/TM-96/R2.

U.S. EPA. 1985. Technical support document for water-quality based toxics control. Office of Water. EPA-440/4-85-032.

U.S. EPA. 1997. Ecological risk assessment guidance for Superfund: Process for designing and conducting ecological risk assessments. Interim Final. EPA 540-R-97-006.

U.S. EPA. 1998. Guidelines For Ecological Risk Assessment. Federal Register Notice. Thursday May 14, 1998. Vol. 63. No. 93. 26846-26924.

U.S. EPA. 2000. Methods for measuring the toxicity and bioaccumulation of sediment-associated contaminants with freshwater invertebrates. Second Edition. EPA/600/R-99/064. Washington, D.C.

U.S. EPA. 2000 (Draft). Equilibrium partitioning sediment guidelines (ESGs) for the protection of benthic organisms: PAH mixtures. Office of Science and Technology and Office of Research and Development.
U.S. EPA 2001. Draft. Equilibrium partitioning sediment guidelines (ESGs) for the protection of benthic organisms: PAH mixtures. Office of Science and Technology and Office of Research and Development.

U.S. EPA. 2002. Memorandum From M. L. Horinko to Superfund National Policy Managers, Regions 1-10. Subject: Principles for Managing Contaminated Sediment Risks at Hazardous Waste Sites. OSWER Directive 9285.6-08.

Appendix Table

1998 and 2001 Bioassay Results Ordered By NOC TPAH Concentrations, Metrics, Benchmarks													
TPAH		2001 Bioassay Results				1998 Bioassay Results					Toxicity Benchmarks		
NOC ug/g OC	Dry Wt. mg/kg	HA-28 UV	HA-28	CT-10	PP-7	PP-7 Elutriate	LV-10	LV-10 UV	DM-48 hr	DM-48 hr UV	EPA ESG TU_{PAHs}	Mac-Donald PEC-Q	% Tests Stat. Sig. Reduc-tions
37SE	0.091					95 - 19 *							0
66 WR	12.6	99	98	85	98						0.8	0.4	0
73 SE	0.18					85 -42*							50
74 W	33	24% LC50											
80 W	40	LC 10											
92 SR	0.42		95	93			132	40*	100	100	1	0.01	14
92 SER						93							
100 W	50			LC 10	LC 10								
115 WR	6.5		96	94			170	0*	92	8*	2	0.2	9
115 WER						83							
120 W	60		LC 10										
143 W	20.1	95	95	93	92						2	0.6	0
146 SE	0.37					78 - 67 *							50
148 SR	3.1	93	86	95	92						2	0.09	0
187 W	90			24% LC50									
228 W	111		24% LC50										
280 W					24% LC50								
292 SE	0.73					38 * - 79*							100
307 W	135	LC50											
399 W	104	36 *									6	3.6	13
584 S	1.5		89 -42 *	80 - 23 *			96 - 61 *	0 *	100	0 *	9		44
584 SE	1.5					5 * - 68 *					9		100
642 S	23	24% LC50											
700 S	25	LC 10		LC 10									
735 S	16.2	99	98	98	96						11	0.6	0
756 S	27			24% LC50									
777 W	373			LC50									
947 W	460		LC50										
1,165 W	570				LC50								
1,310 W	301	33 *	90	88 -30 *	98						19	10.9	25
1,361 WE	23.1					18 * - 66 *							100
1,400 S	50				LC 10								
1,582 W	759	0 *	3 *	5 * - 61 *	18 * - 70 *						23	28	100
1,960 S	70		LC 10										
1,963 W	104	83	90	93	96						28	3.7	0
1,987 S					24% LC50								
2,370 W	94		24% LC50										
2,671 S	85	LC50											
2,722 WE	46.3					0 *							
2,873 W	661	0 *	86	48 * - 40 *	88 - 31 *						42	24	50
3,144 S	111			LC50									
3,996 S	79.9	69 *	95	73	92						62	3	13
4,842 S	823	0 *	46 *	0 *	12 * - 68 *						76	30	100
5,444 WE	92.6					0*							100
8,264 S	296				LC50								
9,861 S	354		LC50										
9,978 S	249	14 *	84	23 *	90 - 33 *						156	9	38
10,888 WE	185					0*							100
21,776 W	370		26 * - 86 *	68 * - 47 *			25 *	0 *	100	0 *	323	13	90
21,776 WE						0 *					323	13	
23,231 S	836	0 *	0 *	0*	0 *						369	32	100
123,56W	235	0 *	0 *	0 *	0 *						1,908	9	100

Notes:

- Column 1: W = Wood Substrate; S = Sand Substrate; E = Elutriate of Sand or Wood Substrate; R = Reference Site
- Columns 3 – 9: Bioassay Results. Top number is % Survival. Bottom negative number is percent reduction in growth at study site compared to reference site. If growth not significantly reduced, no value given. Asterisk indicates statistically significant reduction in survival or growth at study site compared to reference site.
- See text discussion above for derivation of LC 50, 24% LC 50, and LC 10 values in columns 3 – 6.
- See text discussion above or ERA for derivation of the three toxicity benchmarks in columns 12 – 14.

ATTACHMENT 4

Comparison of RI/FS Work Plans From Newfields' 12-15-03 TLR With WDNR Comments

Proposed Sampling Programs

Chequamegon Bay Sediments

Comparison of RI/FS Work Plans From Newfields' 12-15-03 TLR With WDNR Comments - Proposed Sampling Programs
Chequamegon Bay Sediments

	URS Work Plan	SEH Work Plan	Comments
Baseline Problem Formulation	Initial Problem Formulation Process presented in "Strawman Baseline Problem Formulation for Affected Bay Sediments" prepared by URS, March 2003. URS advocates an interactive approach with agencies and other Interested Parties to complete this process. Objectives of the Problem Formulation include:	Presents Section 4.2 'OU#4 Ecological Risk Assessment Problem Formulation and Study Design' in Work Plan.	Have an existing Problem Formulation product based on following process in previous ERAs. Additional iterative study components follow basic PF guidance to design and develop. Present studies in Work Plan under review were the result of responding to CSTAG and other interested party input. No need to do a redundant start-over PF process. Iterative approach has been followed that has resulted in one completed and one planned round of iterative studies based on inputs.
	▪ Develop conceptual site model which incorporates sediment stability evaluation.	▪ Refers to Section 3 'Conceptual Site Model', which describes potential sources, potential exposed receptors, and previous results of analyses that led to preliminary response objectives and remedial action alternatives (SEH, 1998a, b, c in Work Plan references).	CSM from previous ERAs is being used as a basis; is being refined as new information becomes available. Follows the iterative risk assessment process. Sediment stability issue as it relates to potential natural attenuation addressed in past ERAs. URS proposals for sediment stability modeling would be incorporated into existing CSM depending on results of modeling effort. How the model results will be used in management decisions and the time line it will take to produce the model in relation to making timely management decisions for the site needs to be determined.
	▪ Develop risk management goals and objectives.	▪ Management goal stated in Section 4.2.2.1 to 'Restore, protect and maintain the habitat and water quality conditions of the near shore area off Kreher Park…' Ten specific management objectives stated in Section 4.2.2.2.	Refined management goals and objectives in SEH Work Plan. Assume URS is in agreement with goals and objectives as no specific comments made.
	▪ Recommend assessment endpoints and associated risk hypotheses.	▪ Assessment endpoints and risk questions described in Table 16.	Assessment endpoints and risk questions in SEH Work Plans. Assume URS in agreement with endpoints and risk questions as no specific comments made.
	Propose measurement endpoints to address risk hypotheses.	▪ Measurement endpoints described in Table 16.	Measurement endpoints in SEH Work Plan. Assume URS in agreement with endpoints as no specific comments made.
	▪ Use DQO process to develop work scope to address data needs and risk hypotheses, tolerable errors and decision criteria.	▪ Some aspects of the DQO process is described in Section 4.2.3.3, which refers to Tables 17 – 25 for sediment and surface water chemistry as well as some of the validation studies.	DQO process followed. Assume URS agrees with DQO process followed in SEH Work Plan as no specific comments made.
	▪ Sampling stations, sample replication and analytical methodology for any supplemental validation studies will be decided as part of the DQO process during the Problem Formulation.	▪ Section 4.2.3.3.2 describes 'A sample size of 30 would be ideal however due to financial considerations the number of sampling locations will be limited to 8.'	Sample size and sample location needs to be resolved. URS has presented no idea of a sample size or rationale for the size that would be satisfactory to them.

Comparison of RI/FS Work Plans From Newfields' 12-15-03 TLR With WDNR Comments - Proposed Sampling Programs
Chequamegon Bay Sediments

	URS Work Plan	SEH Work Plan	Comments
Baseline Problem Formulation (continued)	▪ Determine range of reference (ambient) concentrations for COPCs. Compare both historical and March 2003 sediment data to sediment data from other nearby areas (e.g., Barksdale) unaffected by point sources of contamination. Perform an evaluation of wood waste impact.	▪ Number of reference stations limited to two. How these will be used to control wood bark, contaminant and grains size differences not specified.	Location of references sites needs to be selected (silty sands and wood dominated). URS indicating determining ranges of ambient concentrations of COPCs implies multiple reference sites rather than one. What sediment data is available from Barksdale site? Bottom characteristics comparable to bay off Kreher Park?
	▪ Analyte list for further sediment and surface water sampling will be determined as part of Problem Formulation	▪ Analysis program to include VOCs, SVOCs, expanded hydrocarbons, TOC, cyanide, copper, lead, mercury, zinc, and AVS/SEM	If URS believes more parameters need to be sampled and analyzed for than are presently listed in the SEH Work Plan, then they need to provide specific parameters now. What information is not available now that would prevent coming up with a specific list? Problem Formulation process to date has not identified additional parameters.
Sediment Characterization Work Proposed	▪ Develop two- and three-dimensional isopleths using both historical and March 2003 sediment data.	▪ Isopleth program unspecified.	NewFields has indicted in their TLR that they have incorporated all existing data into a GIS platform and have produced isopleths of contaminant distribution in the sediments. If satisfactory , no need for SEH Work Plan to duplicate this effort.
	▪ Perform forensic analyses on selected sediment samples to evaluate bioavailablility characteristics.	▪ Specific forensic program for sediments unspecified; however, references to previous methods for analyses will be used if NAPLs encountered.	SEH needs to respond to.
	▪ Perform a sediment stability evaluation. This is described in detail in the URS work plan. This will include evaluation of surface water quality as well as an evaluation of sediment and surface water transport characteristics will be evaluated as part of the sediment stability investigation proposed in the URS work plan.	▪ Sediment stability evaluation unspecified; however, reference in Section 5.5.3 to sediment stability modeling as part of a potential future work plan.	There are no specifics on the extent of surface water or sediment sampling that will be necessary in association with the URS modeling effort.. The SEH Work Plan in Section 5.5.3 states sediment stability modeling may be necessary, would be considered supplemental to the completion of the RI, and will be approved as an individual task if warranted by the agencies. See comments on sediment stability modeling in the above main body of comments.
	▪ One element of this study plan is a comparison of contaminant levels in the water column over undisturbed sediments to levels existing in the water column over recently disturbed sediments. It hasn't been determined whether this evaluation will be after the disturbance resulting from natural events or following manual disturbance.	▪ Surface water column sampling will be performed. Sample locations for ERA to include same eight as above sediment characterization studies. Sample locations for HHRA will include 10 other near shore locations, except one co-located with ERA.	No specific comments made on She designs; assume URS in agreement with designs of surface water sampling effort.

Comparison of RI/FS Work Plans From Newfields' 12-15-03 TLR With WDNR Comments - Proposed Sampling Programs
Chequamegon Bay Sediments

	URS Work Plan	SEH Work Plan	Comments
Biological Studies to Support Baseline Ecological and Human Health Risk Assessment	Other possible studies depending upon outcome of Problem Formulation process:	Studies already proposed in Work Plan based upon SEH Problem Formulation:	URS characterizes the planned studies in the SEH Work Plan as "possible studies". The studies are part of the iterative risk assessment process and are in response to comments from the CSTAG group and other interested parties.
	▪ Further sediment characterization studies may include pore water, benthic community structure, sediment toxicity investigations, caged mussel bioaccumulation studies, fish impact studies, and (potentially) wildlife fish ingestion studies or other studies depending upon the assessment and measurement endpoints agreed upon during the Problem Formulation.	▪ Pore water characterization unspecified; benthic community characterization determined by eight sample locations (two reference locations); toxicity testing program to consist of eight sample locations for 28-day life-cycle chronic toxicity, plus standard and UV light on *Chironomus tentans;* caged mussel program determined by three cages at eight locations (two background) for 90-day in-situ test; fish impact program will consist of eight samples of American smelt to be collected from locations to be determined by WDNR Fisheries staff (eight for HHRA and eight for ERA); same program for two of three species selected from walleye, lake trout and round whitefish; or, yellow perch, smallmouth bass, northern pike or burbot if sufficient samples not retrieved from earlier list, in order of preference; wildlife fish ingestion program unspecified.	URS needs to provide specific comments on the studies in the SEH Work Plan as they are beyond "possible studies" They now are in the realm of planned studies that will be conducted in response to interested party and CSTAG input. As part of the iterative risk assessment process. The only study element that URS has in their list of "possible studies" that SEH's Work Plan does not contain is a pore water study. Pore water characterization is listed as an additional validation study in the August 2003 Draft RI/FS Work Plan, however no specifics are given. See comments in main text above on pore water studies. If URS believes a pore water study is need the specifics and rationales should have been provided.